

PCT

File Copy  
WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification <sup>6</sup> : C12N 15/85, 15/13, C07K 16/28, C12N 15/43, 15/28, C07K 14/525, C12N 15/26, C07K 14/55, C12N 15/19, C07K 14/52, C12N 5/10, 15/62, C07K 19/00, G01N 33/53, 33/60, C12Q 1/68</p>	<p>A1</p>	<p>(11) International Publication Number: <b>WO 98/11241</b>  (43) International Publication Date: 19 March 1998 (19.03.98)</p>
<p>(21) International Application Number: PCT/EP97/04765 (22) International Filing Date: 2 September 1997 (02.09.97) (30) Priority Data: 96114820.2 16 September 1996 (16.09.96) EP (34) Countries for which the regional or international application was filed: GB et al. 96115635.3 30 September 1996 (30.09.96) EP (34) Countries for which the regional or international application was filed: GB et al. (71) Applicant (for all designated States except US): MERCK PATENT GMBH [DE/DE]; Frankfurter Strasse 250, D-64293 Darmstadt (DE). (72) Inventors; and (75) Inventors/Applicants (for US only): VON HOEGEN, Ilka [DE/BE]; Route de Renipont 25a, B-1380 Ohain (BE). BURGE, Christa [DE/DE]; Carsonweg 23, D-64289 Darmstadt (DE). BRÜMMER, Wolfgang [DE/DE]; A, Grenzweg 9, D-64665 Alsbach (DE). DUNKER, Reinhard [DE/DE]; Am Wemsbach 9, D-64354 Reinheim (DE). RIEKE, Erwin</p>	<p>[DE/DE]; Hermannstrasse 12, D-64342 Seeheim-Jugenheim (DE). WELGE, Thomas [DE/DE]; Am Steinernen Kreuz 58a, D-64297 Darmstadt (DE). HAUSER, Hansjörg [DE/DE]; Mascharoder Weg 1, D-38124 Braunschweig (DE). MIELKE, Christian [DE/DE]; Mascharoder Weg 1, D-38124 Braunschweig (DE). (74) Common Representative: MERCK PATENT GMBH; Postfach, D-64271 Darmstadt (DE). (81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i></p>	
<p>(54) Title: OLIGOCISTRONIC EXPRESSION SYSTEM FOR THE PRODUCTION OF HETEROMERIC PROTEINS (57) Abstract  The present invention relates to a mammalian expression system for the production of recombinant heteromeric proteins, preferably antibodies, and more preferably antibody fusion proteins, such as antibody-cytokine fusion proteins, and fragments thereof by means of oligocistrionic expression vectors which are under the control of a strong promoter/enhancer unit, a selection marker gene and at least two IRES elements. The heteromeric fusion proteins can be produced in a robust and stable process in excellent yields.</p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## OLIGOCISTRONIC EXPRESSION SYSTEM FOR THE PRODUCTION OF HETEROMERIC PROTEINS

5 The present invention relates to a mammalian expression system for the production of recombinant heteromeric proteins, preferably antibodies and antibody fusion proteins such as antibody-cytokine fusion proteins and fragments thereof, by means of tri- or oligocistronic expression vectors which are under the control of a strong promoter/enhancer unit and which  
10 contain a selection marker as one of the cistrons. This selection marker guarantees together with at least two IRES elements a robust and stable production of the heteromeric proteins in excellent yields.

### *Background of the invention*

15 For the expression of herteromeric proteins in mammalian cells such as antibody molecules traditionally two vectors have been used which frequently leads to unpredictable overexpression of one of the protein chains in comparison with the second one. Where one chain is relatively overexpressed the cells begin to suffer resulting in instability of production and/or in purification problems (e.g. light chain dimers). One traditional way to overcome this problem is to cotransfer the  
20 vectors in a well defined ratio into the host cells. This requires that the plasmid copies are accepted and integrated simultaneously and stable, and that the plasmid ratio remains constant during cell division. Only for a few systems satisfying results were obtained up to now.

25 Another traditional way is to use independent transcription units located on one plasmid. Thus, the different genes are present on the vector in a correct ratio. Provided that promoters of comparable strength are used equal amounts of the desired protein chains should be obtained. However, different stability and translation efficiencies of the mRNAs which are coding for the different proteins,  
30 and different transcription efficiencies of the genes lead to an unequal synthesis of the desired protein chains.

To avoid these problems di- and multicistronic vectors were developed recently. In such systems the gene units used (coding for the desired proteins, cistrons) are under the control of one single promoter. Normally, only the first cistron located at the 5' terminus is translated efficiently in eucaryotes since the initiation of the translation occurs according to the "cap"- dependent mechanism. The following cistrons are translated insufficiently or not at all. It has been found that the translation of the following cistrons in multicistronic systems can be initiated and pushed by using sequences having no "cap" structure. Such sequences are obtainable from non-translated sections of some viruses, such as poliovirus and encephalomyocarditis virus (Jang et al., 1988, J. Virol. 62:2636; Jang et al., 1989, J. Virol. 63: 1651; Pelletier und Sonnneberg, 1988, Nature 334:320). Within the virus sequences a short section which is not translated and called IRES (internal ribosomal entry site) can be used to allow translational reinitiation independent on the cap. Such sequences have to be interspersed between the cistrons to make a multicistronic mRNA functional. IRES sequences do not influence the "cap"-dependent translation of the first cistron. However, it was found that the "cap" dependent translation is, as a rule, more effective than the IRES-dependent translation which means that the proteins are expressed in a non-stoichiometric ratio and, finally, leads to a loss of stability. Thus, it is very difficult to produce two or more proteins in equimolar ratios even with means of a bi- or oligocistronic expression unit. Biscistronic expression systems and vectors, respectively, using non-antibody genes are known (e.g. Dirks et al., 1993, Gene 128:247). In most of these systems a gene coding for a selection marker was used as second cistron. International patent publication WO 94/05785 discloses a general teaching of expression units in which more than one IRES element can be theoretically inserted into the vector construction. In detail, however, only a bicistronic expression system is described using well defined genes, namely encoding PDGF chains A and B (platelet derived growth factor) separated by an IRES containing unit. No selection marker is used in this system.

It has not been reported until now that heteomeric proteins such as antibody heavy and light chains have been expressed in stoichiometric and stable formation by tri- or oligocistronic systems. It has not been reported, furthermore, that the use of a  
5 selection marker as one of the cistrons leads to transformed cells which have an extraordinary high stability.

Equimolar and stable production of the heteromeric protein chains, such as the heavy and light chain of antibodies, is a prerequisite for a correct association and  
10 folding of the two chains, and, therefore, for a correct steric conformation which is important in order to achieve an optimal biological activity of the associated heteromeric protein or peptide chains.

In the case of an antibody fusion protein, the biologically active ligand for an  
15 antibody-directed targeting should induce the destruction of the target cell either directly or through creating an environment lethal to the target cell. The biologically active ligand can be a cytokine such as IL-1, IL-2, IL-4, IL-6, IL-7, IL-10, IL-13, IFNs, TNF $\alpha$  or CSFs. These cytokines have been shown to elicit anti-tumor effects either directly or by activating host defense  
20 mechanisms (e.g. Mire-Sluis, TIBITECH, 11:74). For instance, IL-2 is considered the central mediator of the immune response. IL-2 has been shown to stimulate the proliferation of T- cells and NK-cells and to induce lymphokine-activated killer cells (LAK). IL-2 enhances the cytotoxicity of T-cells and monocytes. TNF alpha has found a wide application in tumor  
25 therapy, mainly due to its direct cytotoxicity for certain tumor cells and the induction hemorrhagic regression of tumors. In addition TNF alpha potentiates the immune response: it is a costimulant of T-cell proliferation, it induces expression of MHC class I and II antigens and TNF alpha, IFN and IL-1 secretion by macrophages. However, most of the known cytokines  
30 activate effector cells, but show no or only weak chemotactic activity.

5 Chemokines, however, are chemotactic for many effector cells and enhance their presence at the tumor site and induce a variety of effector cell functions (e.g. Miller and Krangel, 1992, "Biology and Biochemistry of the Chemokines,...", Critical Reviews in Immunology 12:17). Examples for suitable chemokines according to the invention are IL-8 and MIP 2 $\alpha$  and MIP 2 $\beta$  which are members of the C-X-C chemokine superfamily (also known as small cytokine superfamily or intecrines).

10 Epidermal growth factor (EGF) is a polypeptide hormone which is mitogenic for epidermal epithelial cells. When EGF interacts with sensitive cells, it binds to membrane receptors (EGFR). The EGFR is a trans-membrane glycoprotein of about 170 kD, and is a gene product of the c-erb-B proto-oncogene.

15 The murine monoclonal antibody mAb425 was raised against the human A431 carcinoma cell line (ATCC CRL 1555; US 5,470,571) and was found to bind to a polypeptide epitope on the external domain of the EGFR. It was found to inhibit the binding of EGF and to mediate tumor cytotoxicity in vitro and to suppress tumor cell growth of epidermal and colorectal carcinoma-derived cell lines in vitro (Rodeck et al., 1987, Cancer Res., 20 47:3692).

25 Humanized and chimeric version of mAb425 are known from WO 92/15683. Fusion proteins of mAb425 (as a whole or fragments thereof) and cytokines or chemokines are described in European patent publications EP 0659 439 and EP 0706 799.

30

**Summary of the invention**

Thus, it is an object of the present invention to provide an expression system  
5 suitable for the stable production of a heteromeric protein, preferably an antibody,  
and more preferably an antibody fusion protein, which avoids the problems of the  
prior art systems as described above.

It has been found as a result of this invention that a proper expression of these  
10 heteromeric proteins can be achieved by using oligocistronic expression units  
comprising at least two IRES elements where the different heteromeric chains, e.g.  
the heavy and light protein chain of an antibody, are cotranslated from one mRNA  
molecule comprising a sequence encoding a selection marker. The strength of the  
effect caused by the selection marker in this system is surprising and could not be  
15 expected compared with usual expression systems of the prior art. The effect is  
especially strong when the gene encoding the selection marker is located at the end  
of all cistrons each separated by IRES units. This is not the case if the selection  
pressure is removed or if the selection marker is used in traditional expression  
vectors. Using the selection marker as last cistron forces the cell to produce the  
20 linked protein / proteins.

The constructs according to the invention allow equimolar production of the  
heteromeric protein chains and guarantee selection and stable, long-term  
expression of the optimal production clones by concomittant expression of the  
25 selection marker, because only those clone will grow under selection pressure  
which express the entire cistronic expression unit.

It has been found that the combination of a selection marker gene and an  
IRES sequence located behind a bicistronic unit (to form a tricistronic unit)  
30 comprising the sequence coding for the light chain of an antibody, an IRES  
sequence and a sequence coding for a fusion protein consisting of the heavy

chain of an antibody fused to another biologically active protein, such as a cytokine or chemokine, is very advantageous with respect to a stable expression in excellent yields.

5

It is an objective of the present invention to provide a new expression system for eucaryotic cells which ensures a stable, reproducible and robust production process for recombinant single and multi-chain protein complexes such as antibodies or, especially, antibody-cytokine fusion proteins.

10

The present invention relates to a mammalian expression system for the production of heteromeric proteins, preferably recombinant antibodies and more preferably antibody fusion proteins such as antibody-cytokine fusion proteins and fragments thereof.

15

The invention relates, preferably, to such a expression system which is able to produce antibody fusion proteins or fragments thereof, wherein the antibody binding sites are directed to the human EGF-receptor and the antibody is covalently linked to a biologically active ligand such as a growth and/or differentiation factor, above all TNF alpha, or IL-2. The invention discloses a set of vectors which comprise oligocistronic, preferably tri- and tetracistronic expression units driven by a single strong promoter hybrid linked to genes encoding protein chains of the light chain, the heavy chain and the active ligand and, additionally a selection marker in the promoter-distal position. Cotranslation of these proteins from one oligocistronic mRNA guarantees strict coupling of expression and allows stoichiometric production of protein chains.

20

25

30



Therefore, it is an object of the invention to provide an oligocistronic expression vector suitable for the production of a heteromeric protein consisting of at least two protein chains in a mammalian host cell comprising

- 5 (i) a promoter / enhancer sequence,
- (ii) a sequence encoding a first chain of the heteromeric protein or a fragment thereof,
- (iii) a sequence encoding a second chain of the heteromeric protein or a fragment thereof,
- 10 (iv) optionally a sequence encoding a third or further chain of the heteromeric protein or a fragment thereof,
- (v) a sequence encoding a selection marker, and
- (vi) at least two sequences comprising a 5'-UTR poliovirus sequence containing an IRES element.

15

It has been found now that the order of the genes located in the vector construct is important with respect to the described advantageous effects. Thus, especially, the gene coding for the selection marker should be located as last cistron within the vector construct. Additionally, in the case of an antibody, the gene encoding the light chain of the antibody should be located in upstream position before the gene coding for the heavy chain.

20

Therefore, it is a preferred object of the invention to provide said expression vector, wherein the sequences (i) to (vi) are in the following order from upstream to downstream progression of said vector construct:

25

- (1) a sequence comprising the promoter / enhancer sequence (i),
- (2) a sequence comprising the sequence encoding a first chain of the heteromeric protein or a fragment thereof (ii),
- (3) a sequence (vi) comprising a first IRES element,
- 30 (4) a sequence comprising the sequence encoding a second chain of the heteromeric protein or a fragment thereof (iii),

- (5) a sequence (vi) comprising a second IRES element,  
(6) optionally a sequence comprising the sequence encoding a third or further chain of the heteromeric protein or a fragment thereof (iv), a sequence comprising  
5 a third or further IRES element (vi) included,  
(7) a sequence comprising the selection marker (v).

The advantage of this system is also shown in Fig. 17 and 18. Under selection  
10 pressure the clones produce in a stable manner the different chains of the heteromeric protein but without selection pressure or "wrong" position of the selection marker the stable productivity is rapidly lost. The greatest advantage of the system is that (heteromeric) proteins can be expressed which can be toxic to the host cells like proteases, glutamate receptor subtypes and serotonin receptor subtypes or antibody fusion proteins wherein the non-antibody partner is normally  
15 highly toxic for the host cells.

Preferably, a corresponding expression system is object of the invention, wherein the sequence (ii) encodes the light chain and the sequence (iii) comprises a sequence encoding the heavy chain of a monoclonal antibody or a fragment thereof.  
20 However, the teaching of this invention is also applicable for heteromeric proteins other than antibodies, for heteromeric proteins having more than two chains, and even normal (one-chain) proteins having toxic activity against the host cell and, finally, heteromeric proteins (e.g. antibody fusion proteins) having strong toxic activity caused by a part of said heteromeric protein..

25 Furthermore, a corresponding expression system is object of the invention, wherein the sequence (iii) consists of two sequences (iiia, iiib), wherein (iiia) encodes the heavy chain of an antibody or a fragment thereof and (iiib) encodes a biologically active ligand, such as a cytokine or a chemokine or a fragment thereof,  
30 in order to form a fusion protein.

5 It has been found, additionally, that such expression vector constructs are preferred, and therefore, object of the invention, wherein the sequence of (iiia) is shortened at its C- terminus and the sequence (iiib) at its N-terminus each by 1 to 15 amino acids.

10 A special and preferred embodiment of the invention is a tricistronic expression vector as defined above and in the claims, wherein the sequence (iiia) and the sequence (iiib) are linked directly in order to encode a fusion protein.

15 In addition the expression vector according to the invention may, optionally, contain eucaryotic sequence elements such as SAR/MAR elements to further increase production and stability of the system. The expression of certain genes has been reported to respond positively to butyrate. The stimulatory effect of butyrate is largest if one or two scaffold/matrix-attached regions (SAR/MAR elements) are present adjacent to the gene (Schlacke et al., 1994, Biochemistry 33:4197). Only after integration of the constructs in to the genome of the host cell these regions increase the expression of adjacent genes in an orientation- and position-independent fashion. Gene activation causes the apparent loss of nucleosome structure ahead of the SAR element and a similar change has been demonstrated by the action of butyrate. Presence of both SARs and butyrate act synergistically in enhancing gene expression (Klehr et al. 1992, Biochemistry 31:3223).

25 Therefore, an expression vector defined above and in the claims is object of the invention, comprising, additionally, one or two, preferably two, SAR elements. Preferably, one SAR element is located in front of the promoter/enhancer region the second one behind the sequence encoding the selection marker. However, other locations are also possible.

30 Preferably, the invention relates to antibody fusion proteins, wherein the non-antibody protein is a biologically active protein. Preferably, such expression

vectors are object of the invention, wherein a sequence (iii**b**) is used which encodes a cytokine or chemokine such as TNF alpha, IL-2 and IL-8.

5 Above all, such expression vectors are object of the invention, wherein the sequences (ii) and (iii) comprise sequences coding for the light and heavy chain of a monoclonal anti-EGFR antibody, preferably, humanized monoclonal antibody 425 (mAb425) or fragments thereof. However, the invention is not restricted to anti-EGFR antibody or mAb425, respectively, but includes also any other  
10 monoclonal antibodies directed to a variety of specificities, for example mAb361.

As an especially preferred embodiment it is object of the invention to provide an expression vector comprising the following units in the given order: the CMV/MPSV promoter/enhancer sequence followed by the sequence encoding the  
15 mAb425 light chain, followed by the sequence from 5'-UTR poliovirus containing an IRES element, followed by a fusion gene encoding a fusion protein consisting of the heavy chain of humanized mAb425 and fused at its C-terminus the sequence encoding TNF alpha or IL-2, followed by another IRES element from 5'- UTR poliovirus, followed by a sequence coding for puromycin acetyl transferase as  
20 selection marker and, finally a nucleotide sequence derived from the polyadenylation signal of SV40.

Furthermore, the well-defined expression vector comprising the nucleotide and amino acid sequences depicted in Figure 15 is object of this invention.  
25

Additionally, it is an object of the invention to provide an expression system comprising a mammalian host cell transformed with an expression vector specified above and in the claims, preferably, wherein the host cell is CHO or BHK.

30 Finally, it is an object of this invention to make available a process for the production of a heteromeric protein, preferably an antibody, especially an antibody

fusion protein, especially a mAb425/TNF alpha or mAb425/IL-2 antibody fusion protein, or fragments thereof, by cultivating the host cells of an expression system as specified above and in the claims in a suitable nutrient and separating, if a  
5 tricistronic vector is used, the complete and active antibody fusion protein from the cells and / or the medium.

#### **Brief Descriptions of the Figures**

##### **10 Fig. 1 (a-e):**

Expression plasmids for the generation of tricistronic expression vectors.

AmpR= Ampicillin resistance gene; IRES = Poliovirus derived internal ribosomal entry site; MPSV = Promoter/Enhancer; CMV = Cytomegalo virus promoter; Puromycin R = Puromycin resistance gene; SV 40 pA = SV  
15 40 polyadenylation site.

##### **Fig. 2:**

Stability of BHK-21 mAb425CH1 clones. Stability of three different clones was determined over the time period indicated. The production of  
20 mAb425CH1 fusion protein of  $10^6$  cells/ml per 24 hrs was determined in an anti-Ig based ELISA. Cells were cultured in medium with (+P) or without (-P) Puromycin.

##### **Fig. 3:**

25 Stability of a BHK21 mAb425CH1-TNF $\alpha$  clone. Cells were cultured in DMEM medium for 89 days without selection pressure. The production of mAb425CH1 fusion protein of  $10^6$  cells/ml per 24 hrs was determined in an anti-Ig based ELISA.

30

**Fig. 4:**

Stability of a BHK21 mAb425CH3-IL-2 clone. Cells were cultured in DMEM medium for 48 days without selection pressure. The production of mAb425CH1 fusion protein of  $10^6$  cells/ml per 24 hrs was determined in an anti-Ig based ELISA.

**Fig. 5:**

SDS PAGE of purified mAb425CH3-IL-2. Lane 1: mAb425CH3-IL-2; Lane 2: IgG1 control antibody. Proteins were run on a 4 to 15% gradient gel (Phast System, Pharmacia) and stained with Coomassie.

**Fig. 6:**

FACS analysis of purified mAb425CH3-IL-2. The human EGF-R-positive carcinoma cell line A431 was incubated with the indicated antibody concentrations. Two different preparations of purified mAb425CH3-IL-2 were compared with purified mAb425 reference antibody.

**Fig. 7:**

Determination of IL-2 activity of purified mAb425CH3-IL-2. IL-2-dependent mouse CTLL2 cells were incubated with mAb425CH3-IL-2 or rec. human IL-2 (WHO Standard). Concentrations of fusion protein are shown as mAb425 equivalents, determined by an ELISA based on an anti-idiotypic antibody specific for mAb425.  $5 \times 10^4$  were cultured for 2 days and pulsed with  $0,5 \mu\text{Ci } ^3\text{H-Thymidine}$  18 hrs before harvesting.

**Fig. 8:**

pMCLDHAP tricistronic vector for the expression of mAb425CH3-TNF $\alpha$ . AmpR= Ampicillin resistance gene; IRES = Poliovirus derived internal ribosomal entry site; MPSV = Promoter/Enhancer; CMV = Cytomegalo

virus promoter; Puromycin R = Puromycin resistance gene; SV 40 pA = SV 40 polyadenylation site.

5 **Fig. 9:**

Stability of a BHK21 mAb425CH3-TNF $\alpha$  clone. Cells were cultured in DMEM medium for 48 days without selection pressure. The production of mAb425CH1 fusion protein of 10<sup>6</sup> cells/ml per 24 hrs was determined in an anti-Ig based ELISA.

10

**Fig. 10:**

Integrity of expression vector DNA in the absence of selective pressure. BHK-21 cell clones transfected with pMCLDHAP and expressing mAb425CH3-TNF $\alpha$  fusion protein were either cultivated under puromycin pressure (+) or grown in the absence of puromycin (-) for the indicated times. Graph A shows antibody fusion protein secretion ( $\mu$ g IgG/ml x 24 hr). B is a Southern blot of chromosomal DNA prepared from cells which were taken at the indicated times. The DNA was restricted with PstI and hybridized with a labelled PstI fragment from pMCLDHAP (1231 bp) encompassing part of the heavy chain fusion protein encoding cDNA (hc). mbh1 represents a single copy DNA fragment (1900 bp) of a hamster c-myc gene which was cohybridized using a specific probe (see example 7). Since both probes are labelled with the same specific activity and their length is similar, the intensity of the hc band corresponds to the copy number of the integrated expression plasmid.

20

25

**Fig. 11:**

FACS analysis of purified mAb425CH3-TNF $\alpha$ . The human EGF-R-positive carcinoma cell line A431 was incubated with the indicated antibody concentrations. Purified mAb425CH3-TNF $\alpha$  was compared with purified humanized mAb425 reference antibody.

30

**Fig. 12:**

Determination of TNF $\alpha$  activity of purified mAb425CH3-TNF $\alpha$  on MCF7 cells. The TNF $\alpha$ -sensitive and EGF-R negative human breast adenocarcinoma cell line MCF7 was used to determine the TNF $\alpha$  activity of the mAb425CH3-TNF $\alpha$  fusion protein. Concentrations of fusion protein are shown as mAb425 equivalents, determined by an ELISA based on an anti-idiotypic antibody specific for mAb425. humanized mAb425 and rTNF $\alpha$  are mixed at a ratio of 6:1 reflecting the molecular ration of both parts in the fusion protein.  $5 \times 10^4$  were cultured for 4 days and pulsed with 0,5  $\mu$ Ci  $^3$ H-Thymidine 18 hrs before harvesting.

**Fig. 13:**

TNF $\alpha$  mediated cytotoxicity of purified mAb425CH3-TNF $\alpha$  is dependent on TNF $\alpha$  sensitivity. The TNF $\alpha$ -resistant and EGF-R-positive human carcinoma cell line A431 was used to determine the specificity of the mAb425CH3-TNF $\alpha$  fusion protein. Concentrations of fusion protein are shown as mAb425 equivalents, determined by an ELISA based on an anti-idiotypic antibody specific for mAb425. Humanized mAb425 and rTNF $\alpha$  are mixed at a ratio of 6:1 reflecting the molecular ratio of both parts in the fusion protein.  $5 \times 10^4$  were cultured for 4 days and pulsed with 0,5  $\mu$ Ci  $^3$ H-Thymidine 18 hrs before harvesting.

**Fig. 14:**

mAb425CH3-TNF $\alpha$  is highly cytotoxic for EGF-R-positive and TNF $\alpha$ -sensitive human tumor cell lines. The human mamma carcinoma cell lines BT20 and the human melanoma cell line C8161 are both TNF $\alpha$ -sensitive and EGF-R-positive. Concentrations of fusion protein are shown as mAb425 equivalents, determined by an ELISA based on an anti-idiotypic antibody specific for mAb425. mAb425 and r TNF $\alpha$  are mixed at a ratio of 6:1 reflecting the molecular ratio of both parts in the fusion protein.  $5 \times 10^4$  were



cultured for 4 days and pulsed with 0,5  $\mu\text{Ci}$   $^3\text{H}$ -Thymidine 18 hrs before harvesting.

5

**Fig. 15:**

Complete nucleotid and amino acid sequence (coding regions) of mAb425CH3-TNF $\alpha$  as shown in Fig. 8.

**Fig. 16:**

10

Hystory of relevant vectors of the invention.

**Fig. 17:**

15

Stability of different antibody fusion protein cell clones (rBHK21mAb425-CH1-IL2). A = mAb425; stability of 3 different clones is tested. The production of fusion protein of  $10^6$  cells / ml in 24 h is determined in the ELISA detecting the antibody part. Cells are cultured for the indicated days in medium with (+P) or without (-P) Puromycin.

**Fig. 18:**

20

Stability of the cell clone rBHK21mAb425-CH3-IL2698-8 with (CHO-M + P) and without (CHO-M - P) selection pressure (puromycin). The stability is tested for 70 days in culture. The production of protein of  $10^6$  cells / ml in 24 h is determined in an ELISA detecting the antibody part of the protein.

25

30

**Detailed Description**

5 Above and below the term "heteromeric protein" means a protein which naturally consists of two or more chains. Only if the corresponding chains are associated and folded correctly the full biological activity of the heteromeric protein can be obtained.

10 Above and below the term "mAb425CH1-" means an antibody construction containing the light chain, the variable region of the heavy chain, and the CH1 domain of the constant region of mAb425.

Above and below the term "mAb425CH2-" means an antibody construction containing the light chain, the variable region of the heavy chain, and the CH1 and CH2 domain of the constant region of mAb425.

15 Above and below the term "mAb425CH3-" means an antibody construction containing the light chain, the variable region of the heavy chain, and the CH1, CH2 and CH3 domain of the constant region of mAb425. This construct corresponds to the complete antibody.

20 Above and below, the term "a sequence encoding ....." does not mean exclusively the specific coding sequence, but may include also a sequence comprising said specific coding sequence, provided that no other statement is made.

25 Said additional sequences indicated above and coding for proteins [ii, iii (iiia, iib), iv, vi] can be prolonged or shortened each by 1 to 20 amino acids provided that the specific biological properties are not substantially amended. Prolongation can be caused, for example, by linker or leader peptides. Furthermore, the expression vector constructs according to the invention may contain introns which are not  
30 translated into amino acids. Prolongations and deletions of coding regions may occur, preferably, at the C- and / or N-terminus of the corresponding specific

peptide or protein. Preferred deletions according to the invention may occur at the C-terminus of the heavy chain of the antibody and the N-terminus of the biological ligand.

5

Furthermore, the invention includes also mutations and variants of the sequences indicated in detail having the same or a very similar biological activity. Such mutations and variants can be produced by accident (e.g. spontaneous mutations, natural radiation) or by intended chemical or physical activities.

10

The term "antibody fragment" means according to the invention an antibody fragment as defined above (mAb-CH1, mAb-CH2) as well as complete antibody (mAb-CH3) which is shortened by 1 to 20 amino acids at the C-terminus of its constant region.

15

The term "biological active ligand" means according to the invention any protein or peptide ligand which is effective against a target cell, above all, against a target cell which is recognized by the antibody part of the antibody fusion protein. The effect of the biological ligand may be, for instance, a toxic and/ or lysing and / or inhibiting one against the target cell, preferably a tumor cell. Examples of suitable biological active ligands are given above.

20

The term "biological active ligand fragment" means according to the present invention a biological ligand (cytokines, chemokines) which is usually shortened by 1 to 20 amino acids at its N-terminus which is connected directly, or optionally via a linker peptide, to the (optionally shortened) C-terminus of the constant region of the antibody heavy chain.

25

All microorganisms, cell lines, plasmids, promoters, resistance markers, replication origins, restriction sites or other fragments or parts of vectors

30

which are mentioned in the description not directly in connection with the claimed invention are commercially or otherwise generally available. Provided that no other hints are given, they are used only as examples and  
5 are not essential with respect to the invention, and can be replaced by other suitable tools and biological materials, respectively.

The techniques which are essential according to the invention are described in detail below and above. Other techniques which are not described in detail  
10 correspond to known standard methods which are well known to a person skilled in the art, or are described mor in detail in the cited references and patent applications and in the standard literature (e.g. Sombrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor, Harlow, Lane, 1988, Antibodies: A Laboratory Manual, Cold Spring  
15 Harbor).

The selection marker according to the invention can be in principal any known selection marker suitable for high expression systems. Examples are enzymes such as puromycin-acetyl transferase or neomycin  
20 phosphotransferase. Puromycin-acetyl transferase is preferred according to this invention.

Alternatively, dominant acting genetic markers useful for monitoring gene transfer in mammalian cells that are based on procaryotic genes encoding  
25 key steps in the synthesis of the essential amino acids, such as tryptophane or histidine can be used. Under appropriate conditions, expression of these genes obviates the nutritional requirements for their respective amino acid products. Expression of the  $\beta$  subunit of tryptophan synthase (trpB, EC 4.2.1.20) of *Escherichia coli* allows mammalian cell survival and  
30 multiplication in medium containing indole in place of tryptophane. The *hisD* gene of *Salmonella typhimurium* encodes histidinol dehydrogenase (EC

1.1.1.23), which catalyses the two-step NAD<sup>+</sup>-dependent oxidation of L-histidinol to L-histidine. In medium lacking histidine and containing histidinol only mammalian cells expressing the hisD gene survive. Use of these markers is advantageous over the use of antibiotics because for either trp or his selection the substitute nutrients indole or histidinol are readily available, inexpensive, stable, permeable to cells and convertible to the end product in a step controlled by one gene (Bode et al. 1995, Int. Rev. Cytol., R. Berezney & K.W. Jeon eds. Academic Press, Vol 162A:389)

As IRES sequences all sequences deriving from viral, synthetic origin or from cells can be used which allow an internal binding of ribosomes. Examples for such sequences are the 5'-UTRs elements from poliovirus type 1, 2 or 3 (picorna virus), from "encephalomyocarditis virus" (EMCV) (Sugimoto et al., 1994, BioTechnol. 12:694), from "Theilers murine encephalomyelitis virus" (TMEV), from "foot and mouth disease virus" (FMDV), from "bovine enterovirus" (BEV), and from "coxsackie B virus" (CBV).

The tri- or oligocistronic expression vector according to the invention works with a single strong promoter/enhancer unit. Examples for suitable promoters/enhancers are: CMV (Boshart et al., 1985, Cell 41:521); MPSV-LTR (Laker et al., 1987, Proc. Natl. Acad. Sci. USA 74,:8458); MPSV-CMV; RSV (Gorman et al., 1982, Proc. Natl. Acad. Sci. USA 79:6777); SV40 (Artelt et al., 1988, Gene 128: 247). The system MPSV(enhancer)-CMV(promoter of the cytomegalie virus) is the preferred unit according to the invention.

The fusion protein described in the examples contains a monoclonal antibody with specificity for the human EGF-receptor(EGFR). The monoclonal mAb425 was raised against the human A431 carcinoma cell line

and found to bind to a polypeptide epitope on the external domain of the EGFR. The heavy chain mAb425 antibody was fused C-terminally to cytokines/chemokines such as IL-2, IL-4, IL-7, TNF $\alpha$  and IL-8 as biologically active ligands. The constructs encoding these immuno-conjugates were generated with recombinant DNA technologies. As pointed out above, the immuno-conjugates contain the variable region of the antibody heavy chain and the CH1 domain of the constant region (antibody-CH1 conjugates), or the CH1 and CH2 domain of the constant region (antibody-CH2 conjugates) or the CH1, CH2 and CH3 domain of the constant region (antibody-CH3 conjugates) fused to the biologically active ligand. By addition of the appropriate light chain immunoconjugates can be generated which target antigen-bearing cells and deliver an active ligand to a specific site in the body. The C-terminal amino acid sequence of the junctional region of CH1 and CH3 fusion proteins is not involved in any secondary structure elements according to the hypothetical computer model. In these regions several putative sites for proteolytic cleavage are present. In order to retain/increase chemical and biological stability these sequences can be shortened up to a limit where the biological activity of the ligand is lost. N-terminal cytokine sequences are frequently involved in receptor binding and biological activity, e.g. in human TNF $\alpha$  amino acid sequences between positions 11 and 35 appear to be critical for receptor binding and triggering of biological responses (Goh & Porter, Prot. Eng. 4:385, 1991). In those cases where loss of activity is caused by inaccessibility of relevant amino acids due to interference of the antibody part linker sequences can be introduced which consist of repetitive units containing amino acids which do not interfere with chemical stability and biological activity, e.g. see Curtis et al. Proc. Natl. Acad. Sci. USA, 88:5809, 1991.

In a preferred embodiment according to the invention a system of expression vectors is provided, which allows easy generation of expression vectors for synthesis of three proteins from a tricistronic expression unit.

5

In a preferred embodiment according to the invention tricistronic vectors have been constructed in which IgG light chain, heavy chain-cytokine fusion protein and a selectable marker are translated from one mRNA. Sequences of translation reinitiation elements (internal ribosomal entry sites = IRES) derived from the 5'-UTR's of poliovirus, which mediate a cap-independent internal initiation of translation, are interspersed between the cistrons.

10

In a preferred embodiment according to the invention the tricistronic mRNA is transcribed from any strong promoter such as a single hybrid MPSV/CMV promoter/enhancer.

15

In a further preferred embodiment the selection marker may be puromycin acetyl transferase, neomycin phosphotransferase or procaryotic genes such as the  $\beta$ -subunit of tryptophane synthase (trpB) derived from *E. coli* or the histidinol dehydrogenase (hisD) of *Salmonella typhimurium* or any resistance marker known in the art. The selection marker is preferably located in the promoter-distal position to ensure stable expression of the entire cistron.

20

In another preferred embodiment of the invention expression is further enhanced by inclusion of one or two, preferably two, scaffold/matrix-attached regions (SAR/MAR elements) into the expression vector. Expression can be synergistically by SAR/MAR elements and butyrate added to the medium.

25

30

In another preferred embodiment of the invention the protein sequence between both parts of the fusion protein can be shortened up to a limit where the biologically active ligand loses its activity.

5

In another preferred embodiment of the invention both parts of the fusion protein can be combined by introducing linker sequences which consist of repetitive units containing preferentially the amino acids alanine, glycine and serine.

10

Furthermore, it is an objective of the invention to manufacture said proteins such as immunoconjugates by transferring the expression vector which contains the tricistronic construct into appropriate host cells such as BHK-21 cells, CHO cells, SP2/0 cells or myeloma cells.

15

Generation of fusion protein constructs consisting of mAb425 and cytokines or chemokines has been disclosed in EP 0659 439 and EP 0706 799, respectively. Fusion proteins have been constructed on the basis of chimeric and humanized mAb425 with cDNAs encoding cytokines such as IL-2, IL-4, IL-7 and TNF $\alpha$  or chemokines such as IL-8 and MIP-2 $\alpha$  and Mip2- $\beta$  fused to the CH1, or CH2 or CH3 domain of the constant region of the mAb425 heavy chain, respectively. The techniques used can be taken, for example from the two European patent publications indicated above which are incorporated in this application by reference.

20

25

The vector system according to the invention leads to a new and innovative production system for high expression of heterodimeric proteins in eucaryotic cells such as antibody-cytokine/chemokine fusion proteins. Light chain and heavy-chain cytokine/chemokine fusion are transcribed together with a selectable marker from one tricistronic mRNA. The advantage of this system is twofold: First, unpredictable overexpression of one of both chains

30



which frequently leads to instability of production and purification problems will be avoided because both chains will be produced at equimolar amounts. Secondly, coupling of product and selection marker in the promoter-distal position guarantees stable and longterm expression of the product. Taken  
5 together, the system described herein represents a robust process for production of complex proteins in eucaryotic cells employing different fermentation techniques.

10 Introduction of vector constructs for the expression of a monovalent immunoconjugate including only the CH1 domain or divalent immunoconjugates including the CH1 and CH2 and CH3 domains into host cells can be achieved by electroporation, DEAE dextrane, calcium phosphate, Lipofectin, protoplast fusion or any known method in the art.

15 Any host cell type may be used provided that the recombinant DNA sequences encoding the immunoconjugate and the appropriate light chain are properly transcribed into mRNA in that cell type. Host cells may be mouse myeloma cells which do not produce immunoglobulin such as Sp2/0-AG14 (ATCC CRL 1581), NSO (Gaffe & Milstein, 1991, Meth. Enzymol.  
20 73(B):3), P3X63Ag8.653 (ATCC CRL 1580) or hamster cells such as CHO-K1 (ATCC CCL 61), or CHO/dhFr- (ATCC CRL 9096), or BHK-21 (ATCC CCL 10). Selection for transfected host cells is done in the presence of the selection marker encoded by the third cistron of the tricistronic expression vector. Clones are analyzed for expression of immunoconjugates by EGF-  
25 receptor or cytokine-specific ELISAs. Selected clones are then further purified by limiting dilution cloning.

30

## **Examples**

### **Example 1**

#### **5      Generation of basic vectors**

The vectors pSBC-1 and pSBC-2 (Dirks et al., 1993, Gene 128:247) have been developed as monocistronic expression vectors. Both vectors contain the SV40 origin of replication, the SV40 early promoter, the SV40 19s splice donor and 19s acceptor, the SV40 polyadenylation signal, procaryotic sequences such as the origin of replication from ColE1 and the Ampicillin resistance gene. In addition pSBC-1 contains the internal ribosomal entry site sequence (IRES) of polio virus for the generation of dicistronic messenger RNAs when appropriately combined with pSBC-2. pSBC vectors were altered by replacing the promoter fragment (ClaI/XhoI) by a hybrid promoter/enhancer composed of an MPSV enhancer of 300 bp (ClaI/XbaI) (Dirks et al., Gene 128:247, 1993) and a PCR amplified huCMVpromoter fragment with XbaI and XhoI ends (bp 220-807 from HEIEE EMBL database) and by replacing the EcoRI-HindIII polylinker by a HindIII-EcoRI polylinker to give pMC-1 (Fig. 1A) and pMC-2 (Fig. 1B), respectively. Based on these vectors a set of vectors have been generated which allow generation of tricistronic expression vectors in a straightforward cloning strategy. The vectors pMC-1 and pMCC-1 (Fig. 1C) are identical except for the multi-cloning sites to facilitate insertion of restriction fragments. In these vectors the promoter-proximal cistron has to be inserted. pMC-2 and pMCC-2 (Fig. 1D) are also identical except for the multi-cloning site and allow expression of one protein chain, but do not contain a selection marker. The vector pMC-2P (Fig. 1E) was created in several steps. First, the blunt-ended fragment of the puromycin resistance gene from pSV2pac (Vara et al. 1986, Nucl. Acid Res. 14:4617) was cloned into the NotI site of pMCC-1. In the resulting plasmid the XbaI/EcoRI was replaced by the analogous fragment from pMCC-2, thereby inserting a new NotI site. The resulting

plasmid is called pMCC-2P (Fig. 1F). pMC-2P was created by exchanging the polylinker into an HindIII/EcoRI polylinker. pMC-2PS (Fig. 1G) was created by insertion of a scaffold-attached region sequence (SAR) of 800 bp from the human Interferon- $\beta$  gene as described (Mielke et al. 1990, Biochemistry 29:7475). All three vectors contain an IRES sequence followed by the selection marker, in this case Puromycin resistance.

After cloning of the respective DNA fragments encoding the protein chains to be expressed into the appropriate vectors generation of a tricistron expression vector is performed as follows: A ClaI/NotI restriction fragment containing the promoter-proximal cistron followed by an IRES sequence is derived from the vectors pMC-1 or pMCC-1, respectively. A NotI/ClaI restriction fragment containing the second cistron followed by an IRES sequence and the selection marker is derived from the vectors pMCC-2P, pMCC-2, pMC-2P, and pMC-2PS. By combination of these two fragments a complete expression vector is generated.

## **Example 2**

### **Cells and gene transfer**

BHK-21 cells (A subclone of ATCC number CCL-10) were cultivated in DMEM supplemented with 10 % fetal calf serum (FCS), 20mM glutamine, 60  $\mu$ g/ml penicillin and 100  $\mu$ g/ml streptomycin.

Calcium phosphate transfections were carried out essentially as described before (Mielke et al. 1990, Biochemistry:29:7474). Minimally 5  $\mu$ g of uncut plasmids were used without the addition of carrier DNA. Stable transfectants were selected and - where indicated - cultivated in medium containing puromycin (Sigma) at a final concentration where only cells expressing the Puromycin resistance marker can grow, e.g. 5  $\mu$ g/ml for BHK-21 cells.

**Example 3****Quantification of secreted antibody**

5  $10^6$  cells/ml were seeded on 25 cm<sup>2</sup> culture flasks in serum free medium and incubated for 24 hours. Medium samples of these cultures were taken for quantification of secreted IgG-chains in a specific ELISA. For this purpose, 96 well immunoplates (Nunc) were coated with an affinity purified goat-anti-human IgG antibody (Fab' specific, Sigma# 1-5260). After incubation with serial dilutions of medium samples, the bound antibody contained in these  
10 samples was detected by application of a peroxidase-conjugated affinity pure goat-anti-human IgG antibody (Dianova#109-035-088) and subsequent staining with ortho-Phenyldiamine-dihydrochloride (OPD)/H<sub>2</sub>O<sub>2</sub>. Quantification was made possible by simultaneous application of an IgG-standard (human IgG1/kappa, Sigma #I3889). No unspecific background was  
15 detectable under these conditions as shown by use of medium supernatants of untransfected cells.

**Example 4****Production of mAb425CH1-IL2 fusion protein****20 Generation of a tricistronic expression vector**

Generation of the DNA sequence encoding the mAb425CH1-IL2 fusion protein has been disclosed in EP 0659 439 and EP 0706 799. A HindIII/EcoRI fragment containing the entire mAb425CH1-IL-2 heavy chain was ligated into the multi-cloning site of the pMC2PSΔH vector. The  
25 NotI/ClaI fragment of this construct was ligated with the ClaI/NotI fragment from pMCLΔHAP containing the mAb425 light chain. The resulting construct contains the light chain in the promoter-proximal position followed by the heavy-chain-IL-2 fusion and the Puromycin resistance. The genes are interspersed by two IRES sequences to allow transcription of all three  
30 cistrons into one messenger RNA.

**Establishment of a recombinant BHK-21 cell line producing mAb425CH1-IL2 fusion protein**

5 BHK-21 (ATCC CCL 10) were transfected with the tricistronic expression vector encoding mAb425CH1-IL2 fusion protein by the calcium phosphate method with a kit commercially available (InVitrogen) according to the manufacturer's instructions. Selection for transfected BHK-21A cells was done in the presence of 5 µg/ml Puromycin (Sigma). Clones are analyzed for expression of immunoconjugates by EGF-receptor or cytokine-specific  
10 ELISAs. Selected clones are then further purified by limiting dilution cloning. In the presence of Puromycin a lot of clones could be isolated which stably express the mAb425CH1-IL2 fusion protein. Three examples are shown in Fig. 2).

15 **Example 5**

**Expression of a mAb425CH1-TNFα fusion protein**

**Generation of a tricistronic expression vector**

Generation of the DNA sequence encoding the mAb425CH1-TNFα fusion protein has been disclosed in EP 0659 439 and EP 0706 799. The heavy  
20 chain-TNFα fusion gene construct was generated on the basis of the heavy chain-IL-2 fusion gene. The KpnI/EcoRI fragment containing part of the heavy chain variable region, the CH1 domain and IL-2 was subcloned into pUC19. In this construct the NcoI/EcoRI fragment containing the IL-2-encoding sequences was exchanged with the NcoI/EcoRI fragment  
25 containing the TNFα-encoding sequences. The KpnI/EcoRI fragment of this construct was combined in pUC18 with the HindIII/KpnI fragment containing the 5'part of the heavy chain variable region to generate the full length heavy chain-TNFα fusion gene. The HindIII/EcoRI fragment was ligated into the multi-cloning site of the pMC2PSΔH vector. The NotI/ClaI  
30 fragment of this construct was ligated with the ClaI/NotI fragment from pMCLΔHAP containing the mAb425 light chain. The resulting construct

contains the light chain in the promoter-proximal position followed by the heavy-chain-TNF $\alpha$  fusion and the Puromycin resistance. The genes are interspersed by two IRES sequences to allow transcription of all three cistrons into one messenger RNA.

**Establishment of a recombinant BHK-21 cell line producing mAb425CH1-TNF $\alpha$  fusion protein**

The establishment of a recombinant BHK-21 cell line producing mAb425CH1-TNF $\alpha$  fusion protein has been performed as described in example 5 for mAb425CH1-IL-2 fusion protein. We could isolate several clones which stably express the mAb425CH1-TNF $\alpha$  fusion protein for more than 12 weeks even without selection pressure. One example is shown in Fig. 3

**Example 6**

**Expression of a mAb425CH3-IL-2 fusion protein**

**Generation of a tricistronic expression vector**

Generation of the DNA sequence encoding the mAb425CH3-IL-2 fusion protein has been disclosed in EP 0659 439 and EP 0706 799. The HindIII/EcoRI fragment containing the complete heavy chain-IL-2 fusion gene was cloned into the multi-cloning vector pMC-2P. The NotI/ClaI fragment of this construct was ligated with the ClaI/NotI fragment from pMCLAHAP containing the mAb425 light chain. The resulting construct contains the light chain in the promoter-proximal position followed by the heavy-chain-TNF $\alpha$  fusion and the Puromycin resistance. The genes are interspersed by two IRES sequences to allow transcription of all three cistrons into one messenger RNA.

**Establishment of a recombinant BHK-21 cell line producing mAb425CH3-IL-2 fusion protein**

Stable BHK-21 cell lines expressing mAb425CH3-IL-2 fusion protein have been established as described in example 5. Several clones could be isolated which stably express the mAb425CH3-IL-2 fusion protein for several weeks even in the absence of selection. One example is shown in Fig. 4

#### **Purification of mAb425CH3-IL2**

Transfected BHK cells (rBHK21A-CH3-IL2/K69-8) were fermented in a semicontinuous manner and the fusion protein was isolated from the collected, cell free supernatant.

The first purification step was performed by affinity chromatography on carrier bound ProteinA (Pharmacia) using the extended bed technology. The starting conditions were 0,1 M phosphate buffer, pH 8,5. Impurities were removed with 0,2 M glycine buffer, pH 5,0 and subsequently, the fusion protein was eluted from the sedimented gel bed with 0,2 M glycine buffer, pH 3,3. The pH of the eluate was immediately neutralized by adding 10 % (vol./vol.) 1 M TRIS solution and brought up to pH 8 - 8,5.

In a second purification step further impurities were separated by cation exchange chromatography on Fractogel EMD  $\text{SO}_3^-$  650(S) (Merck). The starting conditions were 10 mM phosphate buffer, pH 6,0 (conductivity 2 mS). The fusion protein was eluted with a NaCl-gradient 0 - 0,6 M NaCl.

The final purification step was done by size exclusion chromatography on Fractogel BioSEC 650(S) (Merck) in PBS, pH 7,4. Up to 5 % aggregates and small amounts of impurities with smaller molecular weight were separated.

Concentration and diafiltration were done by ultrafiltration (Amicon). Membranes with a cut-off of 30 kDa were used.

Detection of protein-containing fractions was done by SDS-PAGE and an ELISA specific for human Ig with affinity-purified goat anti-human Fc as catcher antibody and affinity-purified goat anti human anti F(ab)<sub>2</sub> coupled to alkaline phosphatase for detection (both Dianova).

The protein content of the preparation was about 1 mg/ml. The endotoxin content was < 1 EU/mg fusion protein. The purity of the protein preparation could be demonstrated by SDS Page (Fig. 5). In Western Blots identity of heavy and light chain could be verified (data not shown).

#### **Functional analysis of recombinant mAb425CH3-IL-2 fusion protein**

FACS analysis with EGF-R-positive cells showed that binding of the antibody portion is identical to a mAb425 control (Fig. 6). Furthermore, IL-2 activity is indistinguishable from the activity of recombinant IL-2 (Fig. 7), indicating that interaction of the fusion protein with the IL-2 receptor is not impaired in the fusion protein. Taken together, it can be concluded that the expression system described herein provides high amounts of the mAb425CH3-IL-2 fusion protein which is fully active with respect to antigen binding and IL-2 activity.

#### **Example 7**

##### **Expression of a mAb425CH3-TNF $\alpha$ fusion protein**

##### **Generation of a tricistronic expression vector**

The PCR amplified coding region of the recombinant light chain (HindIII-EcoRI) gene was inserted into pMC-1 at the polylinker site. The puromycin resistance gene coding sequence was inserted between the IRES sequence and the polyadenylation site of pMC-2 to give pMC-2P. The heavy chain-cytokine fusion protein genes were inserted into the polylinker sequence of pMC-2P. The XmnI/NotI fragments of both Immunoglobulin chain vectors were combined to give e.g. pMCLDHAP, a 8298 bp tricistronic expression vector for IgG-TNF-alpha and puromycin acetyltransferase (Fig. 8).

##### **Establishment of a recombinant BHK-21 cell line producing mAb425CH3-TNF $\alpha$ fusion protein**

BHK-21 cells were transfected with the tricistronic expression vector encoding mAb425CH3-TNF $\alpha$  fusion protein using the calcium phosphate precipitation method as detailed by Mielke et al. (1990, Biochemistry



29:7475). 5 µg of uncut plasmid were used without the addition of carrier DNA. Stable transfectants were selected and cultivated in medium containing Puromycin (Sigma) at a final concentration of 5 µg/ml. Clones are analysed for expression of immunoconjugates by IgG-specific ELISA. Selected clones were further purified by limiting dilution cloning. We could isolate several clones which stably express mAb425CH3-TNFα fusion protein even in the absence of selection. One example is shown in Fig. 9.

#### Chromosomal DNA analysis

Isolation of genomic DNA: Cells from a 141 cm<sup>2</sup> culture dish were harvested in 20 ml TEN buffer [40mM Tris/HCl (pH 7.5), 1mM EDTA, 150 mM NaCl], split into two portions and pelleted for 5 min at 1000 rpm in a table top centrifuge. One of these cell pellets was intensively resuspended in 1 ml of TEN and then provided with 1ml of 2x extraction buffer [20mM tris/HCl (pH 8), 200 mM EDTA, 1 % SDS, 40 µg/ml Rnase A] . After 5 h of incubation at 37 ° C, 50 µl Proteinase K solution (20 mg/ml) was added and incubation was continued over night. Following a standard phenolization step, the DNA solution was dialyzed against TE and was then used without any further precipitation steps.

Southern Blots/Methylation pattern: 20µg of genomic DNA was digested over night with the indicated restriction enzyme in a total volume of 500µl, precipitated by addition of 300 µl 2-propanol and pelleted at 13000 rpm, 4 ° C in a microcentrifuge. DNA pellets were carefully resuspended in 40µl of 1x loading buffer [2.5 % Ficoll (Type 400), 17 mM EDTA, 0.01 % Xylene Cyanol FF), 20µl were applied on a 0.8 % TAE agarose gel and electrophoresed. Gels were then blotted onto nylon membranes (Zeta probe, Biorad) with 0.4 M NaOH over night and membranes were then hybridized to the indicated radiolabelled (Rediprime, Amersham) DNA probes according to manufacturers recommendations and following the protocol of Church and Gilbert (Church, G.M. and Gilbert, W. (1984), PNAS 81, 1991 - 1995). (Fig.10)

**Purification of mAb425CH3-TNF $\alpha$** 

Transformed BHK cells (rBHK21A-CH3-TNF $\alpha$ /SC7.4) were fermented in a semicontinuous manner and the fusion protein was isolated from the collected, cell free supernatant.

The first purification step performed by affinity chromatography on carrier bound ProteinA (Pharmacia) using the extended bed technology. The starting conditions were 0,1 M phosphate buffer, pH 8,5. Impurities were removed with 0,2 M glycine buffer, pH 5,0 before the fusion protein was eluted from the sedimented gel bed with 0,2 M glycine buffer, pH 3,3. The pH of the eluat was immediately brought up to pH 8 - 8,5 by adding 10 % (vol./vol.) 1 M TRIS solution.

The second purification step was done by chromatography on hydroxyapatite (Merck). The starting conditions were 5 mM phosphate, pH 7,0. The elution was performed with a phosphate gradient (5 - 500 mM).

The final purification step was done by size exclusion chromatography on Fractogel BioSEC 650(S) (Merck) in PBS, pH 7,4 as described above. Up to 5 % aggregates and small amounts of impurities with smaller molecular weight were separated.

Concentration and diafiltration were done by ultrafiltration. Membranes with a cut-off of 30 kDa were used.

Detection of protein-containing fractions was done by SDS-PAGE and an ELISA specific for human Ig with affinity-purified goat anti-human Fc as catcher antibody and affinity-purified goat anti human anti F(ab)<sub>2</sub> coupled to alkaline phosphatase for detection (both Dianova).

The protein content of the preparation was about 1 mg/ml. The endotoxin content was < 1 EU/mg fusion protein.

**Assessment of functionality of mAb425CH3-TNF $\alpha$  fusion protein**

The functionality of mAb425CH3-TNF $\alpha$  with respect to antigen binding was demonstrated by FACS analysis (Fig. 11). The fusion protein does bind to EGF-R-positive cells with the same quality as the mAb425 control antibody.

TNF $\alpha$  activity of the mAb425CH3-TNF $\alpha$  fusion protein was investigated on different human tumor cell lines. MCF7 is a human mamma carcinoma cell line which is not EGF-R positive. The inhibition of proliferation is therefore  
5 exclusively based on TNF $\alpha$  activity. As demonstrated in Fig. 12 the growth inhibition induced by the mAb425CH3-TNF $\alpha$  fusion protein is virtually identical to that of recombinant TNF $\alpha$ . mAb425 alone does not have any effect on proliferation of MCF7.

mAb425 was raised against the human carcinoma cell line A431 which is  
10 highly positive for EGF-R expression (Rodeck et al.). It was demonstrated previously that mAb425 is internalized upon binding to A431 cells. A431 is not TNF $\alpha$  sensitive and neither mAb425CH3-TNF $\alpha$  fusion protein nor the combination of mAb425 and recombinant TNF $\alpha$  does have any effect on the growth of A431 cells (Fig. 13) indicating that the growth inhibition  
15 specifically requires expression of TNF $\alpha$  receptors. Lack of TNF $\alpha$  receptors cannot be overcome through internalization of mAb425CH3-TNF $\alpha$  fusion protein mediated by EGF-R receptor.

BT20, a human mamma carcinoma cell line and C8161, a human melanoma cell line, are both EGF-R positive and TNF $\alpha$  sensitive. The density of EGF-R  
20 on the cell surface is higher on BT20 than on C8161 as determined by FACS analysis (data not shown). The proliferation of both cell lines is strongly inhibited by mAb425CH3-TNF $\alpha$  fusion protein (Fig. 14). The effect is more pronounced on BT20 cells than on C8161, which might be due to the increased EGF-R expression which leads to a higher crosslinking of  
25 TNF $\alpha$  receptors and thus increased signal transduction. These experiments clearly demonstrate the superiority of the mAb425CH3-TNF $\alpha$  fusion protein when compared to the combination of mAb425 and TNF $\alpha$ . This could be explained by the crosslinking of TNF $\alpha$  receptors on one side due to capping of EGF-R on the other side. Thereby signal transduction is maximally  
30 enhanced.

SEQUENCE LISTING

## (1) GENERAL INFORMATION:

5

## (i) APPLICANT:

- (A) NAME: Merck Patent GmbH
- (B) STREET: Frankfurter Str. 250
- (C) CITY: Darmstadt
- (E) COUNTRY: Germany
- (F) POSTAL CODE (ZIP): 64271
- (G) TELEPHONE: 49-6151-72-7022
- (H) TELEFAX: 49-6151-72-7191

10

15

- (ii) TITLE OF INVENTION: Oligocistronic Expression System for the Production of Antibody Fusion Proteins

- (iii) NUMBER OF SEQUENCES: 6

20

## (iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

25

## (2) INFORMATION FOR SEQ ID NO: 1:

30

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8298 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

35

- (ii) MOLECULE TYPE: DNA (genomic)

- (iii) HYPOTHETICAL: NO

40

- (iv) ANTI-SENSE: NO

- (v) FRAGMENT TYPE: N-terminal

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: humanized mAb425-TNFalpha Fusion protein
- (B) STRAIN: E. coli K12
- (G) CELL TYPE: Fibroblast
- (H) CELL LINE: BHK-21

45

## (ix) FEATURE:

50

- (A) NAME/KEY: promoter
- (B) LOCATION: 1..904
- (D) OTHER INFORMATION: /function= "Enhancer/promoter: MPSV/CMV"

- 5 (ix) FEATURE:  
    (A) NAME/KEY: intron  
    (B) LOCATION:905..976
- 10 (ix) FEATURE:  
    (A) NAME/KEY: CDS  
    (B) LOCATION:977..1018  
    (D) OTHER INFORMATION:/product= "leader sequence (part)"
- 15 (ix) FEATURE:  
    (A) NAME/KEY: intron  
    (B) LOCATION:1019..1106  
    (D) OTHER INFORMATION:/function= "5'UTR poliovirus"
- 20 (ix) FEATURE:  
    (A) NAME/KEY: CDS  
    (B) LOCATION:1107..1433  
    (D) OTHER INFORMATION:/function= "FRs, CDRs"  
        /product= "light chain hmAb425, variable region,  
                plus leader(rest)"
- 25 (ix) FEATURE:  
    (A) NAME/KEY: intron  
    (B) LOCATION:1434..1595
- 30 (ix) FEATURE:  
    (A) NAME/KEY: CDS  
    (B) LOCATION:1596..1913  
    (D) OTHER INFORMATION:/product= "light chain hmAb425,  
        constant region"
- 35 (ix) FEATURE:  
    (A) NAME/KEY: misc\_feature  
    (B) LOCATION:1914..2581  
    (D) OTHER INFORMATION:/product= "5'UTR from poliovirus +  
        IRES (2029-2159) + intron"
- 40 (ix) FEATURE:  
    (A) NAME/KEY: CDS  
    (B) LOCATION:2582..4537  
    (D) OTHER INFORMATION:/product= "Fusion protein: heavy  
        chain hmAb425 + TNFalpha(from 4064)"
- 45 (ix) FEATURE:  
    (A) NAME/KEY: misc\_RNA  
    (B) LOCATION:4565..5279  
    (D) OTHER INFORMATION:/product= "5'UTR from polivirus  
        plus IRES plus intron"
- 50 (ix) FEATURE:  
    (A) NAME/KEY: CDS  
    (B) LOCATION:5280..5876

(D) OTHER INFORMATION: /function= "selection marker"  
/product= "puromycin acetyl transferase"

## (ix) FEATURE:

5 (A) NAME/KEY: misc\_RNA  
(B) LOCATION: 5877..8298  
(D) OTHER INFORMATION: /product= "DNA sequence comprising  
SV40 PolyA (5929-6181)"  
/standard\_name= "SV40 PolyA"

10

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

15	TCGATAATGA AAGACCCAC CTGTAGGTTT GGCAAGCTAG CTTAAGTAAC GCCATTTTGC	60
	AAGGCATGGG AAAAATACAT AACTGAGAAT AGAGAAGTTC AGATCAAGGT CAGGAACAGA	120
	GAAACAGGAG AATATGGGCC AAACAGGATA TCTGTGGTAA GCAGTTCCTG CCCCCTCAG	180
20	GGCCAAGAAC AGTTGGAACA GGAGAATTGG GCCAAACAGG ATATCTGTGG TAAGCAGTTC	240
	CTGCCCCGCT CAGGGCCAAG AACAGATGGT CCCAGATGC GGTCCCCGCC TCAGCAGTTT	300
	CTAGACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCCGCCATT	360
25	GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA	420
	ATGGGTGGAG TATTTACGGT AAAGTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC	480
30	AAGTACGCCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA	540
	CATGACCTTA TGGGACTTTC CTAATTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC	600
	CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG ACTCAGGGG	660
35	ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG	720
	GGACTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG GTAGGCGTGT	780
40	ACGGTGGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG	840
	CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCGAGGAACT	900
	GGAAAACCAG AAAGTTAACT GGTAAGTTTA GTCTTTTGT CTTTTATTTC AGGTCCCGGA	960
45	ATTAAGCTTC GCCACC ATG GGA TGG AGC TGT ATC ATC CTC TTC TTG GTA	1009
	Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val	
	1 5 10	
50	GCA ACA GCT ACAGGTAAGG GGCTCACAGT AGCAGGCTTG AGGTCTGGAC	1058
	Ala Thr Ala	

	ATATATATGG GTGACAATGA CATCCACTTT GCCTTTCTCT CCACAGGT GTC CAC TCC	1115
	Val His Ser	
	1	
5	GAC ATC CAG ATG ACC CAG AGC CCA AGC AGC CTG AGC GCC AGC GTG GGT	1163
	Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly	
	5 10 15	
10	GAC AGA GTG ACC ATC ACC TGT AGT GCC AGC TCA AGT GTA ACT TAC ATG	1211
	Asp Arg Val Thr Ile Thr Cys Ser Ala Ser Ser Ser Val Thr Tyr Met	
	20 25 30 35	
15	TAT TGG TAC CAG CAG AAG CCA GGT AAG GCT CCA AAG CTG CTG ATC TAC	1259
	Tyr Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr	
	40 45 50	
20	GAC ACA TCC AAC CTG GCT TCT GGT GTG CCA AGC AGA TTC AGC GGT AGC	1307
	Asp Thr Ser Asn Leu Ala Ser Gly Val Pro Ser Arg Phe Ser Gly Ser	
	55 60 65	
25	GGT AGC GGT ACC GAC TAC ACC TTC ACC ATC AGC AGC CTC CAG CCA GAG	1355
	Gly Ser Gly Thr Asp Tyr Thr Phe Thr Ile Ser Ser Leu Gln Pro Glu	
	70 75 80	
30	GAC ATC GCC ACC TAC TAC TGC CAG CAG TGG AGT AGT CAC ATA TTC ACG	1403
	Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Trp Ser Ser His Ile Phe Thr	
	85 90 95	
35	TTC GGC CAA GGG ACC AAG GTG GAA ATC AAA CGTGAGTAGA ATTTAACTT	1453
	Phe Gly Gln Gly Thr Lys Val Glu Ile Lys	
	100 105	
40	TGCTTCCTCA GTTGGATCCA TCTGGGATAA GCATGCTGTT TTCTGTCTGT CCCTAACATG	1513
45	CCCTGTGATT ATGCGCAAAC AACACACCCA AGGGCAGAAC TTTGTTACTT AAACACCATC	1573
50	CTGTTTGCTT CTTTCCTCAG GA ACT GTG GCT GCA CCA TCT GTC TTC ATC TTC	1625
	Thr Val Ala Ala Pro Ser Val Phe Ile Phe	
	1 5 10	
55	CCG CCA TCT GAT GAG CAG TTG AAA TCT GGA ACT GCC TCT GTT GTG TGC	1673
	Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys	
	15 20 25	
60	CTG CTG AAT AAC TTC TAT CCC AGA GAG GCC AAA GTA CAG TGG AAG GTG	1721
	Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val	
	30 35 40	
65	GAT AAC GCC CTC CAA TCG GGT AAC TCC CAG GAG AGT GTC ACA GAG CAG	1769
	Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln	
	45 50 55	

	GAC AGC AAG GAC AGC ACC TAC AGC CTC AGC AGC ACC CTG ACG CTG AGC Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser	1817
	60 65 70	
5	AAA GCA GAC TAC GAG AAA CAC AAA GTC TAC GCC TGC GAA GTC ACC CAT Lys Ala Asp Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His	1865
	75 80 85 90	
10	CAG GGC CTG AGC TCG CCC GTC ACA AAG AGC TTC AAC AGG GGA GAG TGT Gln Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys	1913
	95 100 105	
	TAGAATTCAG CTTTAAAC AGCTCTGGGG TTGTACCCAC CCCAGAGGCC CACGTGGCGG	1973
15	CTAGTACTCC GGTATTGCGG TACCCTTGTA CGCCTGTTTT ATACTCCCTT CCCGTAACCT	2033
	AGACGCACAA AACCAAGTTC AATAGAAGGG GGTACAAACC AGTACCACCA CGAACAAGCA	2093
	CTTCTGTTTC CCCGGTGATG TCGTATAGAC TGCTTGCGTG GTTGAAAGCG ACGGATCCGT	2153
20	TATCCGCTTA TGTACTTCGA GAAGCCCACT ACCACCTCGG AATCTTCGAT GCGTTGCGCT	2213
	CAGCACTCAA CCCCAGAGTG TAGCTTAGGC TGATGAGTCT GGACATCCCT CACCGGTGAC	2273
25	GGTGGTCCAG GCTGCGTTGG CGGCCTACCT ATGGCTAACG CCATGGGACG CTAGTTGTGA	2333
	ACAAGGTGTG AAGAGCCTAT TGAGCTACAT AAGAATCCTC CGGCCCTGA ATGCGGCTAA	2393
	TCCCAACCTC GGAGCAGGTG GTCACAAACC AGTGATTGGC CTGTCGTAAC GCGCAAGTCC	2453
30	GTGGCGSAAC CGACTACTTT GGGTGTCCGT GTTTCCTTTT ATTTTATTGT GGCTGCTTAT	2513
	GGTGACAATC ACAGATTGTT ATCATAAAGC GAATTGGATT GCGGCCGCGA ATTAAGCTTC	2573
35	CCGCCACC ATG GAC TGG ACC TGG CGC GTG TTT TGC CTG CTC GCC GTG GCT Met Asp Trp Thr Trp Arg Val Phe Cys Leu Leu Ala Val Ala	2623
	1 5 10	
40	CCT GGG GCC CAC AGC CAG GTG CAA CTA GTG CAG TCC GGC GCC GAA GTG Pro Gly Ala His Ser Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val	2671
	15 20 25 30	
	AAG AAA CCC GGT GCT TCC GTG AAG GTG AGC TGT AAA GCT AGC GGT TAT Lys Lys Pro Gly Ala Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr	2719
45	35 40 45	
	ACC TTC ACA TCC CAC TGG ATG CAT TGG GTT AGA CAG GCC CCA GGC CAA Thr Phe Thr Ser His Trp Met His Trp Val Arg Gln Ala Pro Gly Gln	2767
	50 55 60	
50	GGG CTC GAG TGG ATT GGC GAG TTC AAC CCT TCA AAT GGC CGG ACA AAT Gly Leu Glu Trp Ile Gly Glu Phe Asn Pro Ser Asn Gly Arg Thr Asn	2815
	65 70 75	



	TAT AAC GAG AAG TTT AAG AGC AAG GCT ACC ATG ACC GTG GAC ACC TCT	2863
	Tyr Asn Glu Lys Phe Lys Ser Lys Ala Thr Met Thr Val Asp Thr Ser	
	80 85 90	
5	ACA AAC ACC GCC TAC ATG GAA CTG TCC AGC CTG CGC TCC GAG GAC ACT	2911
	Thr Asn Thr Ala Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr	
	95 100 105 110	
10	GCA GTC TAC TAC TGC GCC TCA CGG GAT TAC GAT TAC GAT GGC AGA TAC	2959
	Ala Val Tyr Tyr Cys Ala Ser Arg Asp Tyr Asp Tyr Asp Gly Arg Tyr	
	115 120 125	
15	TTC GAC TAT TGG GGA CAG GGT ACC CTT GTC ACC GTC AGT TCA GGT GAG	3007
	Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Gly Glu	
	130 135 140	
20	TGG ATC CTC TGC GCC TGG GCC CAG CTC TGT CCC ACA CCG CGG TCA CAT	3055
	Trp Ile Leu Cys Ala Trp Ala Gln Leu Cys Pro Thr Pro Arg Ser His	
	145 150 155	
25	GGC ACC ACC TCT CTT GCA GCC TCC ACC AAG GGC CCA TCG GTC TTC CCC	3103
	Gly Thr Thr Ser Leu Ala Ala Ser Thr Lys Gly Pro Ser Val Phe Pro	
	160 165 170	
30	CTG GCA CCC TCC TCC AAG AGC ACC TCT GGG GGC ACA GCG GCC CTG GGC	3151
	Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly	
	175 180 185 190	
35	TGC CTG GTC AAG GAC TAC TTC CCC GAA CCG GTG ACG GTG TCG TGG AAC	3199
	Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn	
	195 200 205	
40	TCA GGC GCC CTG ACC AGC GGC GTG CAC ACC TTC CCG GCT GTC CTA CAG	3247
	Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln	
	210 215 220	
45	TCC TCA GGA CTC TAC TCC CTC AGC AGC GTG GTG ACC GTG CCC TCC AGC	3295
	Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser	
	225 230 235	
50	AGC TTG GGC ACC CAG ACC TAC ATC TGC AAC GTG AAT CAC AAG CCC AGC	3343
	Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser	
	240 245 250	
55	AAC ACC AAG GTG GAC AAG AAA GTT GAG CCC AAA TCT TGT GAC AAA ACT	3391
	Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr	
	255 260 265 270	
60	CAC ACA TGC CCA CCG TGC CCA GCA CCT GAA CTC CTG GGG GGA CCG TCA	3439
	His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser	
	275 280 285	

	GTC TTC CTC TTC CCC CCA AAA CCC AAG GAC ACC CTC ATG ATC TCC CGG	3487
	Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg	
	290 295 300	
5	ACC CCT GAG GTC ACA TGC GTG GTG GTG GAC GTG AGC CAC GAA GAC CCT	3535
	Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro	
	305 310 315	
10	GAG GTC AAG TTC AAC TGG TAC GTG GAC GGC GTG GAG GTG CAT AAT GCC	3583
	Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala	
	320 325 330	
15	AAG ACA AAG CCG CGG GAG GAG CAG TAC AAC AGC ACG TAC CGG GTG GTC	3631
	Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val	
	335 340 345 350	
20	AGC GTC CTC ACC GTC CTG CAC CAG GAC TGG CTG AAT GGC AAG GAG TAC	3679
	Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr	
	355 360 365	
	AAG TGC AAG GTC TCC AAC AAA GCC CTC CCA GCC CCC ATC GAG AAA ACC	3727
	Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr	
	370 375 380	
25	ATC TCC AAA GCC AAA GGG CAG CCC CGA GAA CCA CAG GTG TAC ACC CTG	3775
	Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu	
	385 390 395	
30	CCC CCA TCC CGG GAT GAG CTG ACC AAG AAC CAG GTC AGC CTG ACC TGC	3823
	Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys	
	400 405 410	
35	CTG GTC AAA GGC TTC TAT CCC AGC GAC ATC GCC GTG GAG TGG GAG AGC	3871
	Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser	
	415 420 425 430	
40	AAT GGG CAG CCG GAG AAC AAC TAC AAG ACC ACG CCT CCC GTG CTG GAC	3919
	Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp	
	435 440 445	
	TCC GAC GGC TCC TTC TTC CTC TAC AGC AAG CTC ACC GTG GAC AAG AGC	3967
	Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser	
	450 455 460	
45	AGG TGG CAG CAG GGG AAC GTC TTC TCA TGC TCC GTG ATG CAT GAG GCT	4015
	Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala	
	465 470 475	
50	CTG CAC AAC CAC TAC ACG CAG AAG AGC CTC TCC CTG TCT CCG GGT AAA	4063
	Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys	
	480 485 490	

	ATG GTC AGA TCA TCT TCG CGA ACC CCG AGT GAC AAG CCT GTA GCC CAT	4111
	Met Val Arg Ser Ser Ser Arg Thr Pro Ser Asp Lys Pro Val Ala His	
	495 500 505 510	
5	GTT GTA GCA AAC CCT CAA GCT GAG GGG CAA CTG CAG TGG CTG AAC CGC	4159
	Val Val Ala Asn Pro Gln Ala Glu Gly Gln Leu Gln Trp Leu Asn Arg	
	515 520 525	
10	CGG GCC AAT GCC CTC CTG GCC AAT GGC GTC GAG CTG AGA GAT AAC CAG	4207
	Arg Ala Asn Ala Leu Leu Ala Asn Gly Val Glu Leu Arg Asp Asn Gln	
	530 535 540	
15	CTG GTG GTG CCA TCA GAG GGC CTG TAC CTC ATC TAC TCC CAG GTC CTC	4255
	Leu Val Val Pro Ser Glu Gly Leu Tyr Leu Ile Tyr Ser Gln Val Leu	
	545 550 555	
20	TTC AAG GGC CAA GGC TGC CCG TCG ACC CAT GTG CTC CTC ACC CAC ACC	4303
	Phe Lys Gly Gln Gly Cys Pro Ser Thr His Val Leu Leu Thr His Thr	
	560 565 570	
25	ATC AGC CGC ATC GCC GTC TCC TAC CAG ACC AAG GTT AAC CTC CTC TCT	4351
	Ile Ser Arg Ile Ala Val Ser Tyr Gln Thr Lys Val Asn Leu Leu Ser	
	575 580 585 590	
30	GCC ATC AAG AGC CCC TGC CAG AGG GAG ACC CCA GAG GGG GCT GAG GCC	4399
	Ala Ile Lys Ser Pro Cys Gln Arg Glu Thr Pro Glu Gly Ala Glu Ala	
	595 600 605	
35	AAG CCC TGG TAT GAG CCC ATC TAT CTG GGA GGG GTC TTC CAG CTC GAG	4447
	Lys Pro Trp Tyr Glu Pro Ile Tyr Leu Gly Gly Val Phe Gln Leu Glu	
	610 615 620	
40	AAG GGT GAC CGA CTC AGC GCT GAG ATC AAT CGG CCC GAC TAT CTC GAC	4495
	Lys Gly Asp Arg Leu Ser Ala Glu Ile Asn Arg Pro Asp Tyr Leu Asp	
	625 630 635	
45	TTT GCC GAG TCC GGA CAG GTC TAC TTT GGG ATC ATT GCC CTG	4537
	Phe Ala Glu Ser Gly Gln Val Tyr Phe Gly Ile Ile Ala Leu	
	640 645 650	
50	TGATAAGGAT CCCCAGGTAC CGAGCTCGAA TTCAGCTTTT AAAACAGCTC TGGGGTTGTA	4597
	CCCCACCCAG AGGCCACGT GCGGGCTAGT ACTCCGGTAT TCGGGTACCC TTGTACGCT	4657
	GTTTTATACT CCCTTCCCGT AACTTAGACG CACAAAACCA AGTTCAATAG AAGGGGGTAC	4717
	AAACCACTAC CACCACGAAC AAGCACTTCT GTTCCCCGG TGATGTCGTA TAGACTGCTT	4777
	GCGTGGTTGA AAGCGACGGA TCCGTTATCC GCTTATGTAC TTCGAGAAGC CCAGTACCAC	4837
	CTCGGAATCT TCGATGCGTT GCGCTCAGCA CTCAACCCCA GAGTGTAGCT TAGGCTGATG	4897
	AGTCTGGACA TCCCTCACCG GTGACGGTGG TCCAGGCTGC GTTGGCGGCC TACCTATGGC	4957

	TAACGCCATG GGACGCTAGT TGTGAACAAG GTGTGAAGAG CCTATTGAGC TACATAAGAA	5017
5	TCCTCCGGCC CCTGAATGCG GCTAATCCCA ACCTCGGAGC AGGTGGTCAC AAACCASTGA	5077
	TTGGCCTGTC GTAACGCGCA AGTCCGTGGC GGAACCGACT ACTTTGGGTG TCCGTGTTTC	5137
	CTTTTATTTT ATTGTGGCTG CTTATGGTGA CAATCACAGA TTGTTATCAT AAAGCGAATT	5197
10	GGATTGCGGC CGGCCGCCAC GACCGGTGCC GCCACCATCC CCTGACCCAC GCCCCTGACC	5257
	CCTCACAAGG AGACGACCTT CC ATG ACC GAG TAC AAG CCC ACG GTG CGC CTC	5309
	Met Thr Glu Tyr Lys Pro Thr Val Arg Leu	
15	1 5 10	
	GCC ACC CGC GAC GAC GTC CCC CGG GCC GTA CGC ACC CTC GCC GCC GCG	5357
	Ala Thr Arg Asp Asp Val Pro Arg Ala Val Arg Thr Leu Ala Ala Ala	
	15 20 25	
20	TTC GCC GAC TAC CCC GCC ACG CGC CAC ACC GTC GAC CCG GAC CGC CAC	5405
	Phe Ala Asp Tyr Pro Ala Thr Arg His Thr Val Asp Pro Asp Arg His	
	30 35 40	
25	ATC GAG CGG GTC ACC GAG CTG CAA GAA CTC TTC CTC ACG CGC GTC GGG	5453
	Ile Glu Arg Val Thr Glu Leu Gln Glu Leu Phe Leu Thr Arg Val Gly	
	45 50 55	
30	CTC GAC ATC GGC AAG GTG TGG GTC GCG GAC GAC GGC GCC GCG GTG GCG	5501
	Leu Asp Ile Gly Lys Val Trp Val Ala Asp Asp Gly Ala Ala Val Ala	
	60 65 70	
35	GTC TGG ACC ACG CCG GAG AGC GTC GAA GCG GGG GCG GTG TTC GCC GAG	5549
	Val Trp Thr Thr Pro Glu Ser Val Glu Ala Gly Ala Val Phe Ala Glu	
	75 80 85 90	
	ATC GGC CCG CGC ATG GCC GAG TTG AGC GGT TCC CGG CTG GCC GCG CAG	5597
	Ile Gly Pro Arg Met Ala Glu Leu Ser Gly Ser Arg Leu Ala Ala Gln	
	95 100 105	
40	CAA CAG ATG GAA GGC CTC CTG GCG CCG CAC CGG CCC AAG GAG CCC GCG	5645
	Gln Gln Met Glu Gly Leu Leu Ala Pro His Arg Pro Lys Glu Pro Ala	
	110 115 120	
45	TGG TTC CTG GCC ACC GTC GGC GTC TCG CCC GAC CAC CAG GGC AAG GGT	5693
	Trp Phe Leu Ala Thr Val Gly Val Ser Pro Asp His Gln Gly Lys Gly	
	125 130 135	
50	CTG GGC AGC GCC GTC GTG CTC CCC GGA GTG GAG GCG GCC GAG CGC GCC	5741
	Leu Gly Ser Ala Val Val Leu Pro Gly Val Glu Ala Ala Glu Arg Ala	
	140 145 150	

	GGG GTG CCC GCC TTC CTG GAG ACC TCC GCG CCC CGC AAC CTC CCC TTC	5789
	Gly Val Pro Ala Phe Leu Glu Thr Ser Ala Pro Arg Asn Leu Pro Phe	
	155 160 165 170	
5	TAC GAG CGG CTC GGC TTC ACC GTC ACC GCC GAC GTC GAG TGC CCG AAG	5837
	Tyr Glu Arg Leu Gly Phe Thr Val Thr Ala Asp Val Glu Cys Pro Lys	
	175 180 185	
10	GAC CGC GCG ACC TGG TGC ATG ACC CGC AAG CCC GGT GCC TGACGCCCCG	5886
	Asp Arg Ala Thr Trp Cys Met Thr Arg Lys Pro Gly Ala	
	190 195	
	CCCACGACCC GCAGCGCCCG ACCGAAAGGA GCGCACGACC CCATGAGCTT CGATCCAGAC	5946
15	ATGATAAGAT ACATTGATGA GTTTGGACAA ACCACAATA GAATGCAGTG AAAAAAATGC	6006
	TTTATTTGTG AAATTTGTGA TGCTATTGCT TTATTTGTAA CCATTATAAG CTGCAATAAA	6066
	CAAGTTAACA ACAACAATTG CATTCATTTT ATGTTTCAGG TTCAGGGGGA GGTGTGGGAG	6126
20	GTTTTTTAAA GCAAGTAAAA CCTCTACAAA TGTGGTATGG CTGATTATGA TCCTGCCTCG	6186
	CGCGTTTCGG TGATGACGGT GAAACCTCT GACACATGCA GCTCCCGGAG ACGGTCACAG	6246
25	CTTGCTCTGA AGCGGATGCC GGGAGCAGAC AAGCCCGTCA GGGCGCGTCA GCGGGTGTG	6306
	GCGGGTGTG GGGCGCAGCC ATGACCCAGT CACGTAGCGA TAGCGGAGTG TATACTGGCT	6366
	TAACATATGCG GCATCAGAGC AGATTGTACT GAGAGTGCAC CATATGTCGG GCCGCGTTGC	6426
30	TGGCGTTTTT CCATAGGCTC CGCCCCCTG ACGAGCATCA CAAAAATCGA CGCTCAAGTC	6486
	AGAGGTGGCG AAACCCGACA GGAATAAAA GATACCAGGC GTTTCCCCCT GGAAGCTCCC	6546
35	TCGTGCGCTC TCCTGTTCCG ACCCTGCCGC TTACCGGATA CCTGTCCGCC TTTCTCCCTT	6606
	CGGGAAGCGT GGCCTTTCT CATAGCTCAC GCTGTAGGTA TCTCAGTTCG GTGTAGGTCG	6666
	TTGCTCCAA GCTGGGCTGT GTGCACGAAC CCCCCTTCA GCCCGACCGC TGCGCCTTAT	6726
40	CCGGTAACTA TCGTCTTGAG TCCAACCCGG TAAGACACGA CTTATCGCCA CTGGCAGCAG	6786
	CCACTGGTAA CAGGATTAGC AGAGCGAGGT ATGTAGGCGG TGCTACAGAG TTCTTGAAGT	6846
45	GGTGGCCTAA CTACGGCTAC ACTAGAAGGA CAGTATTTGG TATCTGCGCT CTGCTGAAGC	6906
	CAGTTACCTT CGGAAAAAGA GTTGGTAGCT CTTGATCCGG CAAACAAACC ACCGCTGGTA	6966
	GCGGTGGTTT TTTTGTGTC AAGCAGCAGA TTACGCGCAG AAAAAAGGA TCTCAAGAAG	7026
50	ATCCTTTGAT CTTTCTACG GGGTCTGACG CTCAGTGGAA CGAAAACTCA CGTTAAGGGA	7086
	TTTTGGTCAT GAGATTATCA AAAAGGATCT TCACCTAGAT CCTTTTAAAT TAAAAATGAA	7146

	GTTTTAAATC AATCTAAAGT ATATATGAGT AAACCTGGTC TGACAGTTAC CAATGCTTAA	7206
5	TCAGTGAGGC ACCTATCTCA GCGATCTGTC TATTTCTGTTT ATCCATAGTT GCCTGACTCC	7266
	CCGTCGTGTA GATAACTACG ATACGGGAGG GCTTACCATC TGGCCCCAGT GCTGCAATGA	7326
	TACCGCGAGA CCCACGCTCA CCGGCTCCAG ATTTATCAGC AATAAACCAG CCAGCCGGAA	7386
10	GGGCCGAGCG CAGAAGTGGT CCTGCAACTT TATCCGCCTC CATCCAGTCT ATTAATTGTT	7446
	GCCGGGAAGC TAGAGTAAGT AGTTCGCCAG TTAATAGTGC GCAACGTTGT TGCCATTGCT	7506
15	ACAGGCATCG TGGTGTACG CTCGTCGTTT GGTATGGCTT CATTTCAGCTC CGGTTCCCAA	7566
	CGATCAAGGC GAGTTACATG ATCCCCATG TTGTGCAAAA AAGCGGTTAG CTCCTTCGGT	7626
	CCTCCGATCG TTGTCAGAAG TAAGTTGGCC GCAGTGTAT CACTCATGST TATGGCAGCA	7686
20	CTGCATAATT CTCTTACTGT CATGCCATCC GTAAGATGCT TTTCTGTGAC TGGTGASTAC	7746
	TCAACCAAGT CATTCTGAGA ATAGTGTATG CGGCGACCGA GTTGCTCTTG CCCGGCGTCA	7806
25	ACACGGGATA ATACCGCGCC ACATAGCAGA ACTTTAAAAG TGCTCATCAT TGGAAAACGT	7866
	TCTTCGGGGC GAAAACTCTC AAGGATCTTA CCGCTGTTGA GATCCAGTTC GATGTAACCC	7926
	ACTCGTGCAC CCAACTGATC TTCAGCATCT TTTACTTTCA CCAGCGTTTC TGGGTGAGCA	7986
30	AAAACAGGAA GGCAAAATGC CGCAAAAAG GGAATAAGGG CGACACGGAA ATGTTGAATA	8046
	CTCATACTCT TCCTTTTTCA ATATTATTGA AGCATTTATC AGGGTTATTG TCTCATGAGC	8106
35	GGATACATAT TTGAATGTAT TTAGAAAAAT AAACAAATAG GGGTTCCGCG CACATTTCCT	8166
	CGAAAAGTGC CACCTGACGT CTAAGAAACC ATTATTATCA TGACATTAACT CTATAAAAAT	8226
	AGGCGTATCA CGAGGCCCTT TCGTCTTCAA GAATTGGTCG ATCGACCAAT TCTCATGTTT	8286
40	GACAGCTTAT CA	8298

## (2) INFORMATION FOR SEQ ID NO: 2:

- 45 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 14 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear
- 50 (ii) MOLECULE TYPE: protein  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala

1 5 10

(2) INFORMATION FOR SEQ ID NO: 3:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 109 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Val His Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala  
15 1 5 10 15

Ser Val Gly Asp Arg Val Thr Ile Thr Cys Ser Ala Ser Ser Ser Val  
20 25 30

Thr Tyr Met Tyr Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu  
35 40 45

Leu Ile Tyr Asp Thr Ser Asn Leu Ala Ser Gly Val Pro Ser Arg Phe  
50 55 60

Ser Gly Ser Gly Ser Gly Thr Asp Tyr Thr Phe Thr Ile Ser Ser Leu  
65 70 75 80

Gln Pro Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Trp Ser Ser His  
85 90 95

Ile Phe Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys  
100 105

35

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 106 amino acids

40

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

45

Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln  
1 5 10 15

Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr  
20 25 30

Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser  
35 40 45

Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr  
 50 55 60  
 5 Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys  
 65 70 75 80  
 His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro  
 85 90 95  
 10 Val Thr Lys Ser Phe Asn Arg Gly Glu Cys  
 100 105

15 (2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 652 amino acids

(B) TYPE: amino acid

20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

25 Met Asp Trp Thr Trp Arg Val Phe Cys Leu Leu Ala Val Ala Pro Gly  
 1 5 10 15  
 Ala His Ser Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys  
 20 25 30  
 30 Pro Gly Ala Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe  
 35 40 45  
 Thr Ser His Trp Met His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu  
 35 50 55 60  
 Glu Trp Ile Gly Glu Phe Asn Pro Ser Asn Gly Arg Thr Asn Tyr Asn  
 65 70 75 80  
 40 Glu Lys Phe Lys Ser Lys Ala Thr Met Thr Val Asp Thr Ser Thr Asn  
 85 90 95  
 Thr Ala Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val  
 100 105 110  
 45 Tyr Tyr Cys Ala Ser Arg Asp Tyr Asp Tyr Asp Gly Arg Tyr Phe Asp  
 115 120 125  
 Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Gly Glu Trp Ile  
 50 130 135 140  
 Leu Cys Ala Trp Ala Gln Leu Cys Pro Thr Pro Arg Ser His Gly Thr  
 145 150 155 160



Thr Ser Leu Ala Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala  
 165 170 175  
 5 Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu  
 180 185 190  
 Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly  
 195 200 205  
 10 Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser  
 210 215 220  
 Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu  
 15 225 230 235 240  
 Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr  
 245 250 255  
 20 Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr  
 260 265 270  
 Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe  
 275 280 285  
 25 Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro  
 290 295 300  
 Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val  
 30 305 310 315 320  
 Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr  
 325 330 335  
 35 Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val  
 340 345 350  
 Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys  
 355 360 365  
 40 Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser  
 370 375 380  
 Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro  
 45 385 390 395 400  
 Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val  
 405 410 415  
 50 Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly  
 420 425 430

Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp  
 435 440 445  
 Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp  
 5 450 455 460  
 Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His  
 465 470 475 480  
 10 Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys Met Val  
 485 490 495  
 Arg Ser Ser Ser Arg Thr Pro Ser Asp Lys Pro Val Ala His Val Val  
 15 500 505 510  
 Ala Asn Pro Gln Ala Glu Gly Gln Leu Gln Trp Leu Asn Arg Arg Ala  
 515 520 525  
 20 Asn Ala Leu Leu Ala Asn Gly Val Glu Leu Arg Asp Asn Gln Leu Val  
 530 535 540  
 Val Pro Ser Glu Gly Leu Tyr Leu Ile Tyr Ser Gln Val Leu Phe Lys  
 545 550 555 560  
 25 Gly Gln Gly Cys Pro Ser Thr His Val Leu Leu Thr His Thr Ile Ser  
 565 570 575  
 Arg Ile Ala Val Ser Tyr Gln Thr Lys Val Asn Leu Leu Ser Ala Ile  
 30 580 585 590  
 Lys Ser Pro Cys Gln Arg Glu Thr Pro Glu Gly Ala Glu Ala Lys Pro  
 595 600 605  
 35 Trp Tyr Glu Pro Ile Tyr Leu Gly Gly Val Phe Gln Leu Glu Lys Gly  
 610 615 620  
 Asp Arg Leu Ser Ala Glu Ile Asn Arg Pro Asp Tyr Leu Asp Phe Ala  
 625 630 635 640  
 40 Glu Ser Gly Gln Val Tyr Phe Gly Ile Ile Ala Leu  
 645 650

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 199 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr Arg Asp Asp Val  
 1 5 10 15  
 Pro Arg Ala Val Arg Thr Leu Ala Ala Ala Phe Ala Asp Tyr Pro Ala  
 5 20 25 30  
 Thr Arg His Thr Val Asp Pro Asp Arg His Ile Glu Arg Val Thr Glu  
 35 40 45  
 10 Leu Gln Glu Leu Phe Leu Thr Arg Val Gly Leu Asp Ile Gly Lys Val  
 50 55 60  
 Trp Val Ala Asp Asp Gly Ala Ala Val Ala Val Trp Thr Thr Pro Glu  
 65 70 75 80  
 15 Ser Val Glu Ala Gly Ala Val Phe Ala Glu Ile Gly Pro Arg Met Ala  
 85 90 95  
 Glu Leu Ser Gly Ser Arg Leu Ala Ala Gln Gln Gln Met Glu Gly Leu  
 20 100 105 110  
 Leu Ala Pro His Arg Pro Lys Glu Pro Ala Trp Phe Leu Ala Thr Val  
 115 120 125  
 25 Gly Val Ser Pro Asp His Gln Gly Lys Gly Leu Gly Ser Ala Val Val  
 130 135 140  
 Leu Pro Gly Val Glu Ala Ala Glu Arg Ala Gly Val Pro Ala Phe Leu  
 145 150 155 160  
 30 Glu Thr Ser Ala Pro Arg Asn Leu Pro Phe Tyr Glu Arg Leu Gly Phe  
 165 170 175  
 Thr Val Thr Ala Asp Val Glu Cys Pro Lys Asp Arg Ala Thr Trp Cys  
 35 180 185 190  
 Met Thr Arg Lys Pro Gly Ala  
 195

## Patent Claims

- 5 1. Oligocistronic expression vector suitable for the production of a heteromeric protein consisting of at least two protein chains in a mammalian host cell comprising
- (i) a promoter / enhancer sequence,
  - (ii) a sequence encoding a first chain of the heteromeric protein or a fragment thereof,
  - 10 (iii) a sequence encoding a second chain of the heteromeric protein or a fragment thereof,
  - (iv) optionally a sequence encoding a third or further chain of the heteromeric protein or a fragment thereof,
  - (v) a sequence encoding a selection marker, and
  - 15 (vi) at least two sequences comprising a 5'-UTR poliovirus sequence containing an IRES element.
2. Expression vector according to claim 1, wherein the sequences (i) to (vi) are in the following order from upstream to downstream progression of said vector construct:
- 20 (1) a sequence comprising the promoter / enhancer sequence (i),
  - (2) sequence comprising the sequence encoding a first chain of the heteromeric protein or a fragment thereof (ii),
  - (3) a sequence (vi) comprising a first IRES element,
  - 25 (4) a sequence comprising the sequence encoding a second chain of the heteromeric protein or a fragment thereof (iii),
  - (5) a sequence (vi) comprising a second IRES element,
  - (6) optionally a sequence comprising the sequence encoding a third or chain of the heteromeric protein or a fragment thereof (iv), and a
  - 30 sequence comprising a third or further IRES element (vi) located behind

the third or further sequence encoding the corresponding chain,  
(7) a sequence comprising the selection marker (v).

- 5        3. Tricistronic expression vector according to claim 1 or 2 (comprising two IRES elements) wherein the sequence (ii) encodes the light chain and the sequence (iii) comprises a sequence encoding the heavy chain of a monoclonal antibody (iiia), and sequences (iv) are not present.
- 10       4. Tricistronic expression vector according to claim 3, wherein the sequence (iii) comprises besides sequence (iiia) a sequence (iiib) encoding a biologically active ligand in order to produce an antibody fusion protein.
- 15       5. Expression vector according to claims 3 to 4 wherein the sequence (iiia) is shortened at its C-terminus and the sequence (iiib) is shortened at its N-terminus by a number of nucleotides each coding for 1 to 20 amino acids.
- 20       6. Expression vector according to claims 3 to 5, wherein a sequence (iiib) is used encoding a cytokine or chemokine.
- 25       7. Expression vector according to claim 6, wherein a sequence (iiib) is used encoding TNF alpha or IL-2.
- 30       8. Expression vector according to claim 1 to 7, wherein sequences (ii) and (iii) encoding the light and heavy chain of a monoclonal anti-EGFR antibody are used.
9. Expression vector according to claim 8 comprising the sequences encoding humanized monoclonal antibody 425 (mAb425).

10. Expression vector according to claim 3 comprising the CMV/MPSV promoter/enhancer sequence followed by the sequence encoding the mAb425 light chain, followed by the sequence from 5' UTR poliovirus containing an IRES element, followed by a fusion gene encoding a fusion protein consisting of the heavy chain of humanized mAb425 and fused at its C-terminus the sequence encoding TNF alpha or IL-2, followed by another IRES element from 5' UTR poliovirus, followed by a sequence coding for puromycin acetyl transferase and, finally the sequence of the polyadenylation signal of SV40.
11. Expression vector according to claim 10 comprising the DNA sequence which codes for the amino acid each depicted in Fig. 15.
12. Expression vector according to claims 1 to 10, comprising, additionally, two SAR elements.
13. Expression system comprising a mammalian host cell transformed with an expression vector specified in one of the claims 1 to 12.
14. Expression system according to claim 13, wherein the host cell is CHO, BHK-21 or SP2/0.
15. Process for the production of a heteromeric protein or fragments thereof by cultivating the host cells of an expression system specified in claim 13 in a suitable nutrient and separating the complete and active heteromeric protein from the cells and / or the medium.
16. Process according to claim 15 for the production of mAb425/TNF-alpha or mAb425/IL-2 Antibody fusion proteins or fragments thereof.

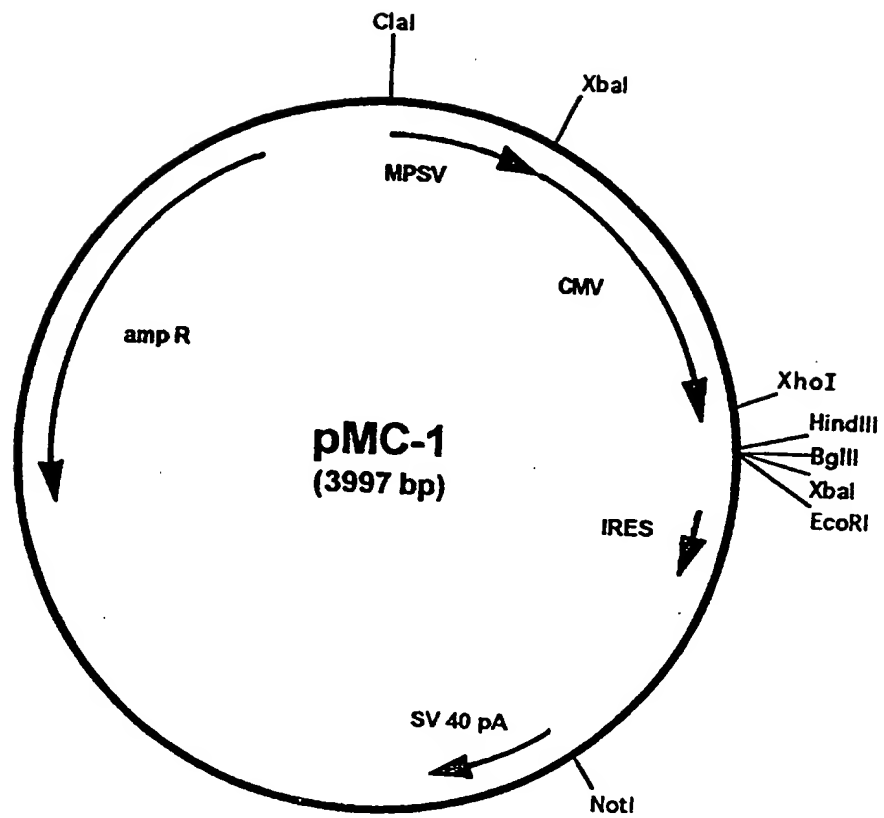


FIG. 1 A

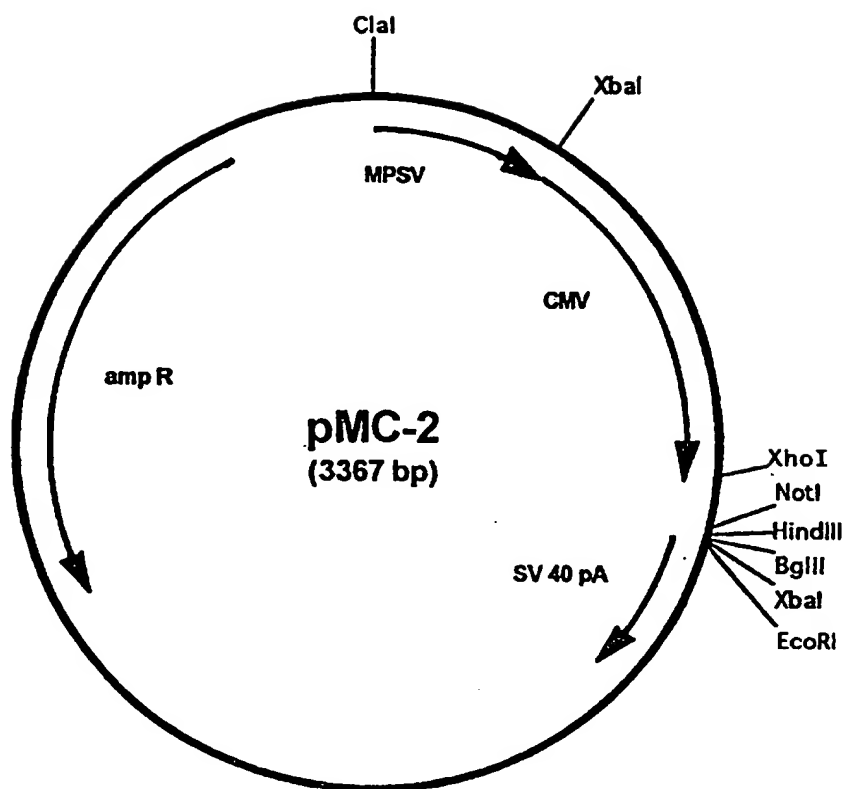


FIG. 1 B



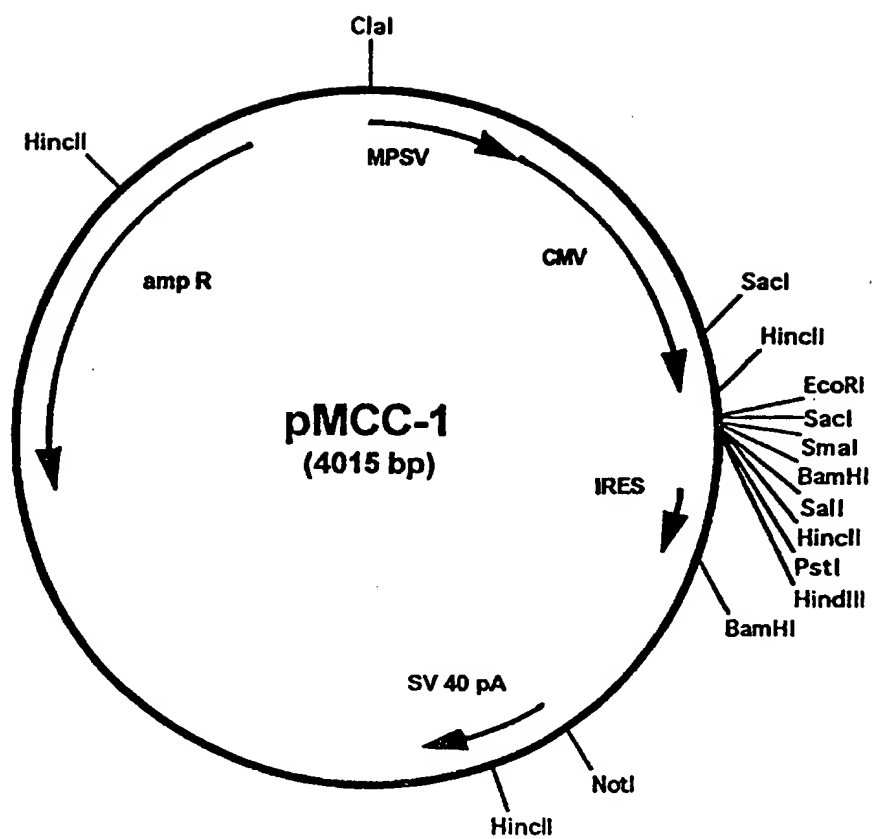


FIG. 1 C

4/31

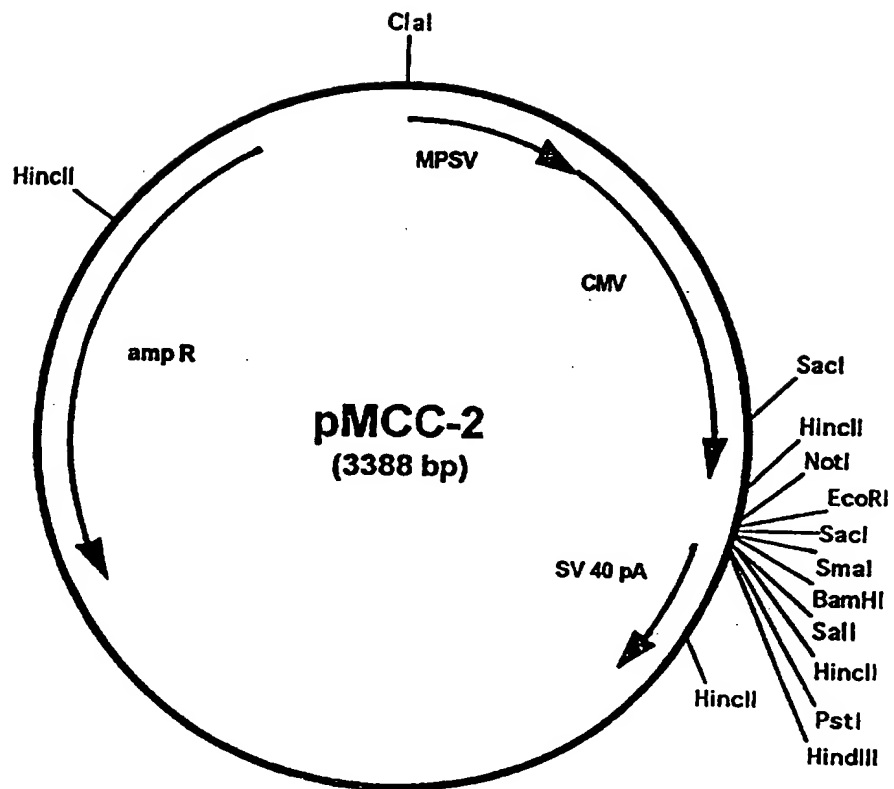


FIG. 1 D

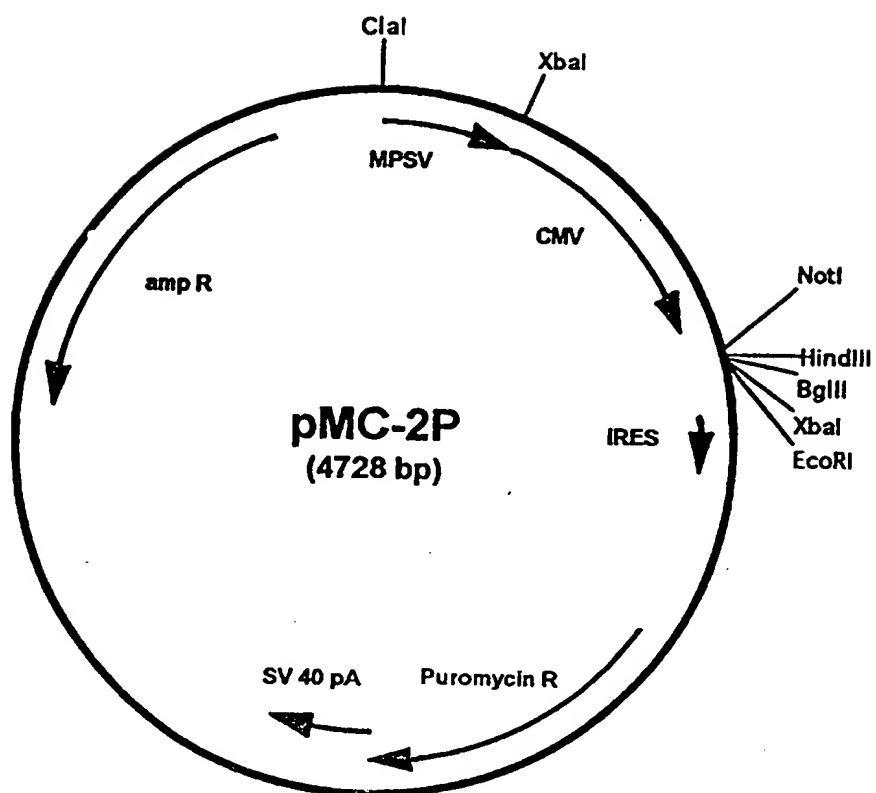


FIG. 1 E

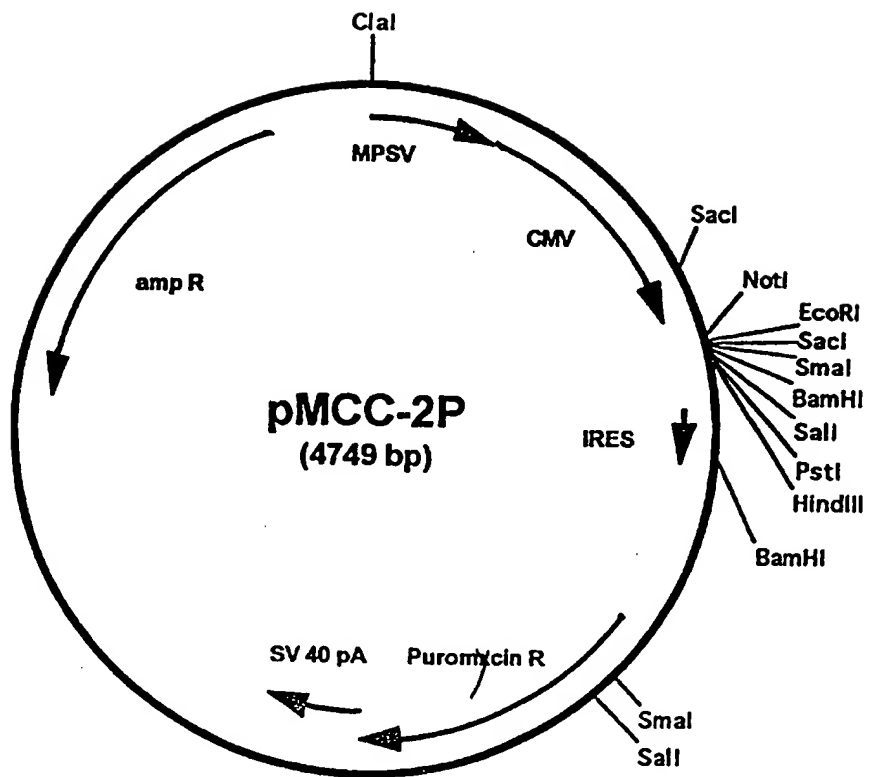


FIG. 1 F

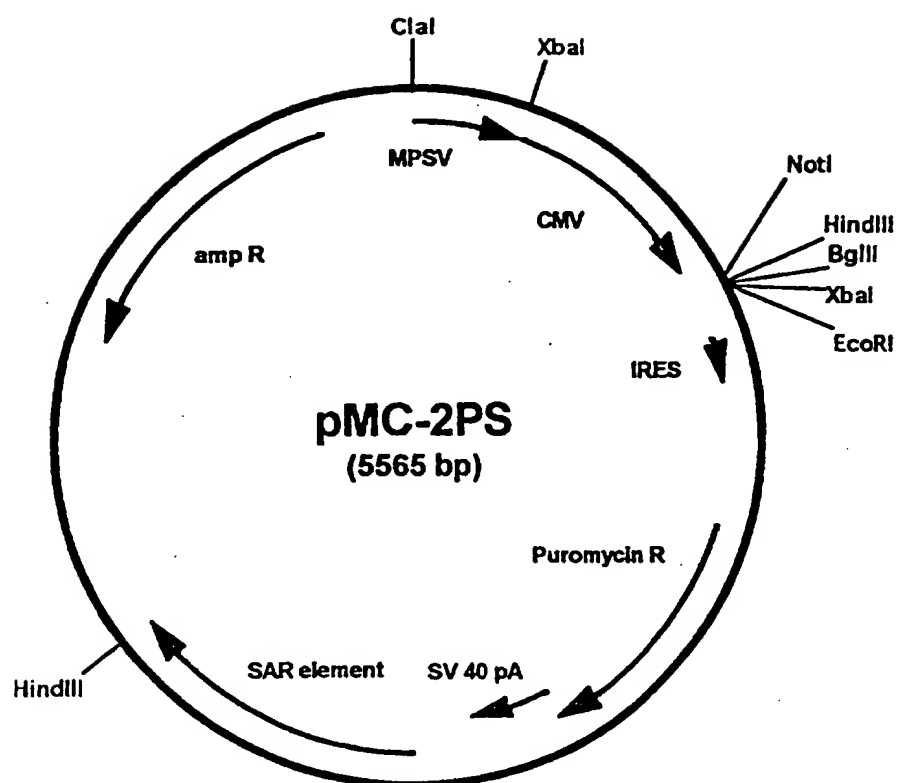


FIG. 1 G

8/31

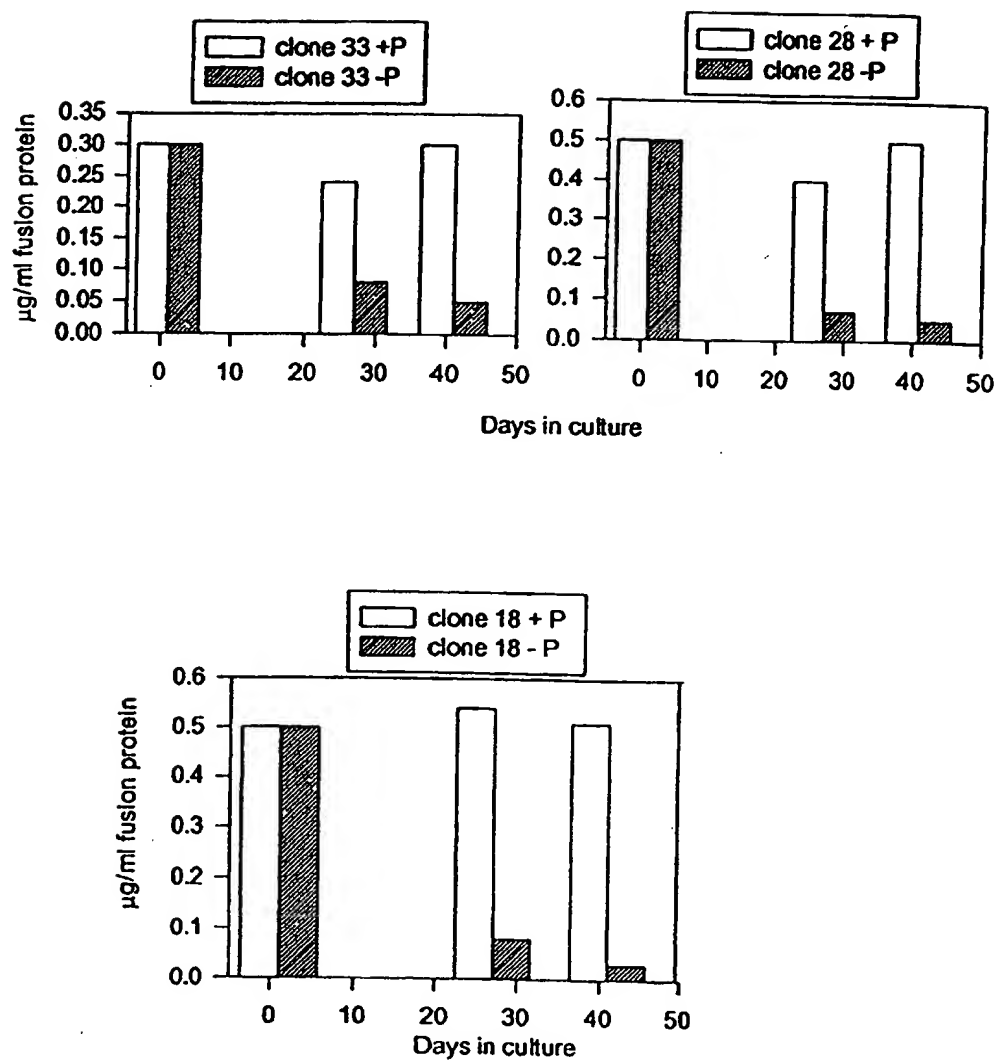


FIG. 2

9/31

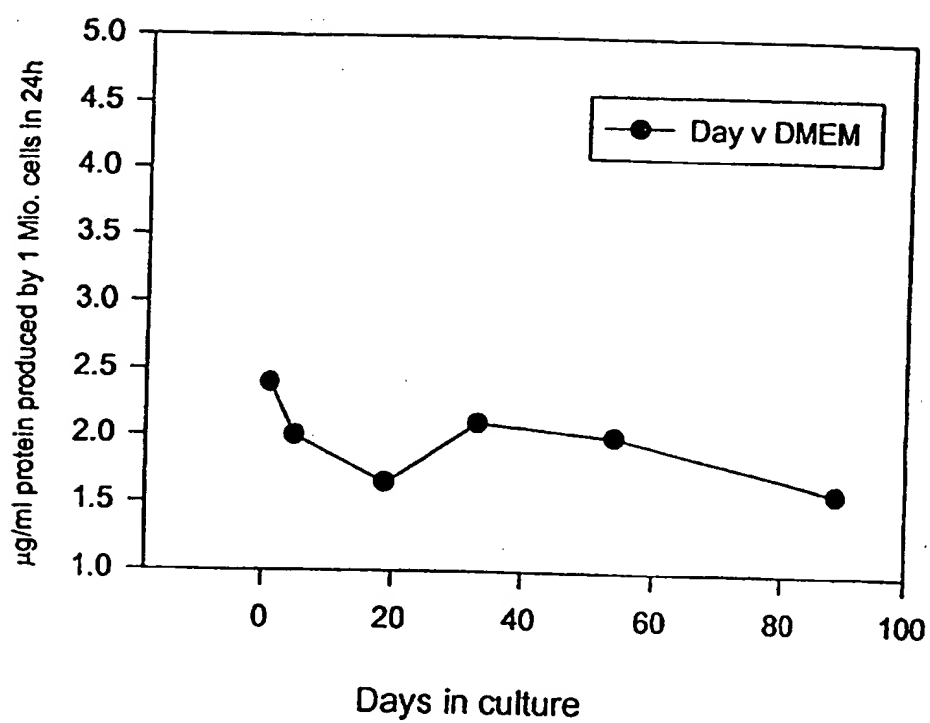


FIG. 3

10/31

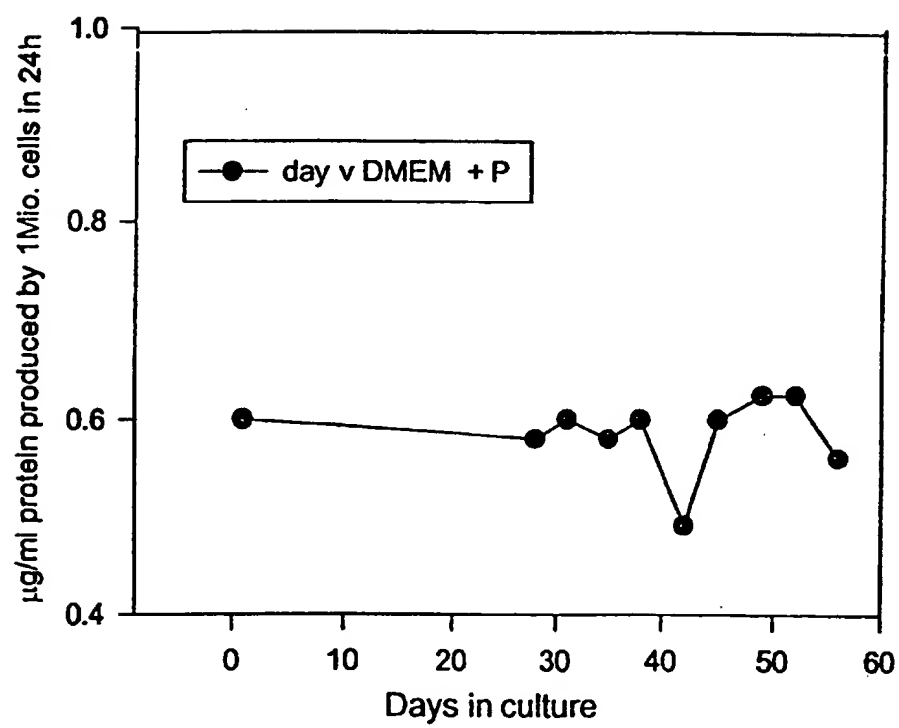


FIG. 4



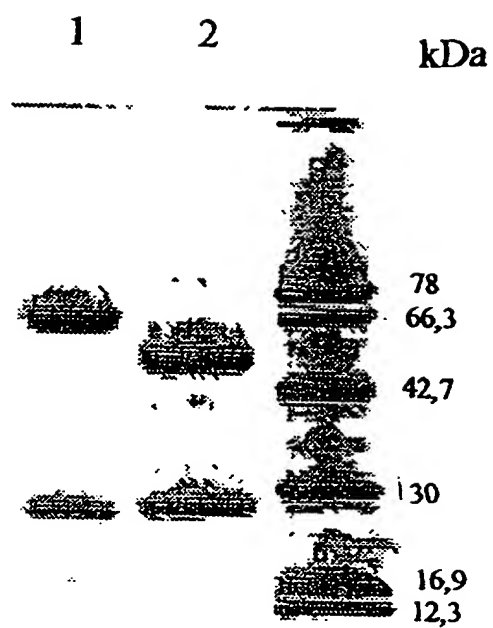


FIG. 5

12/31

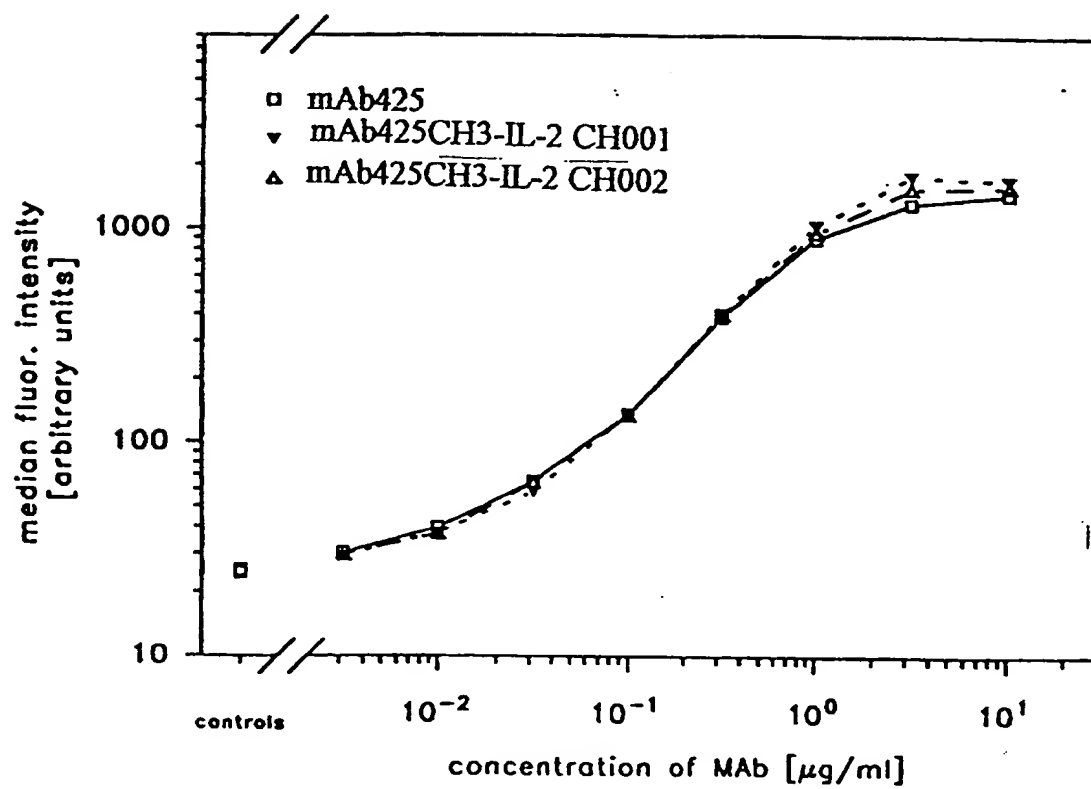


FIG. 6

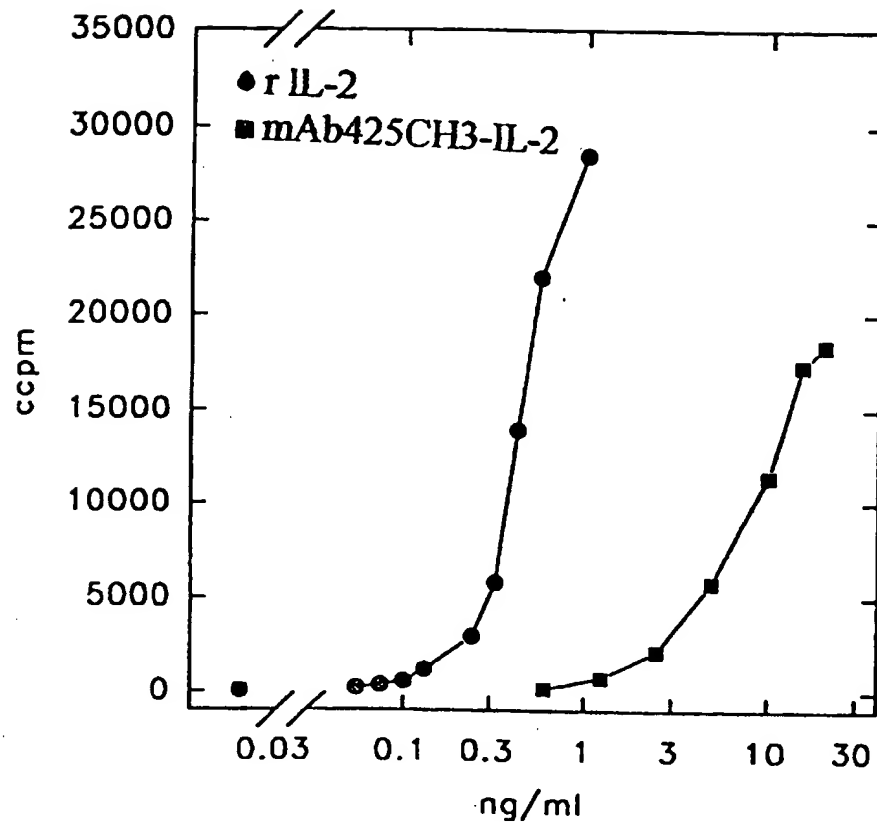


FIG. 7

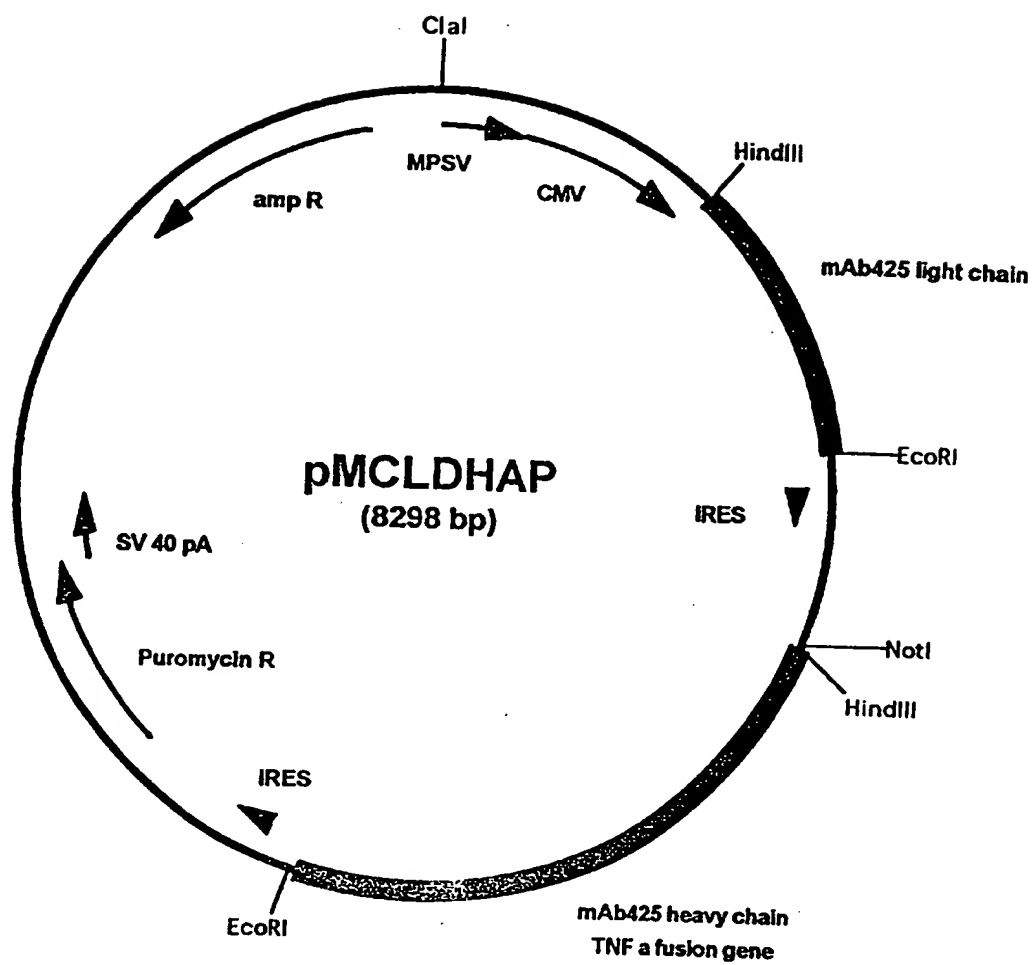


FIG. 8

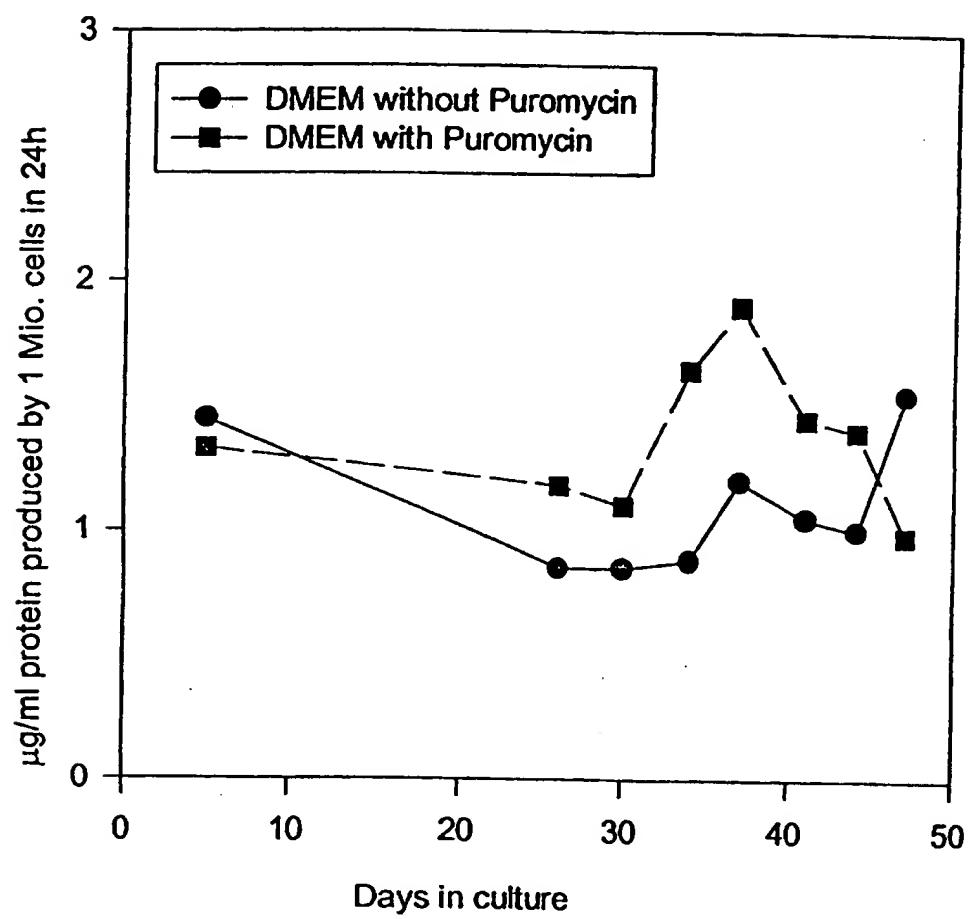


FIG. 9

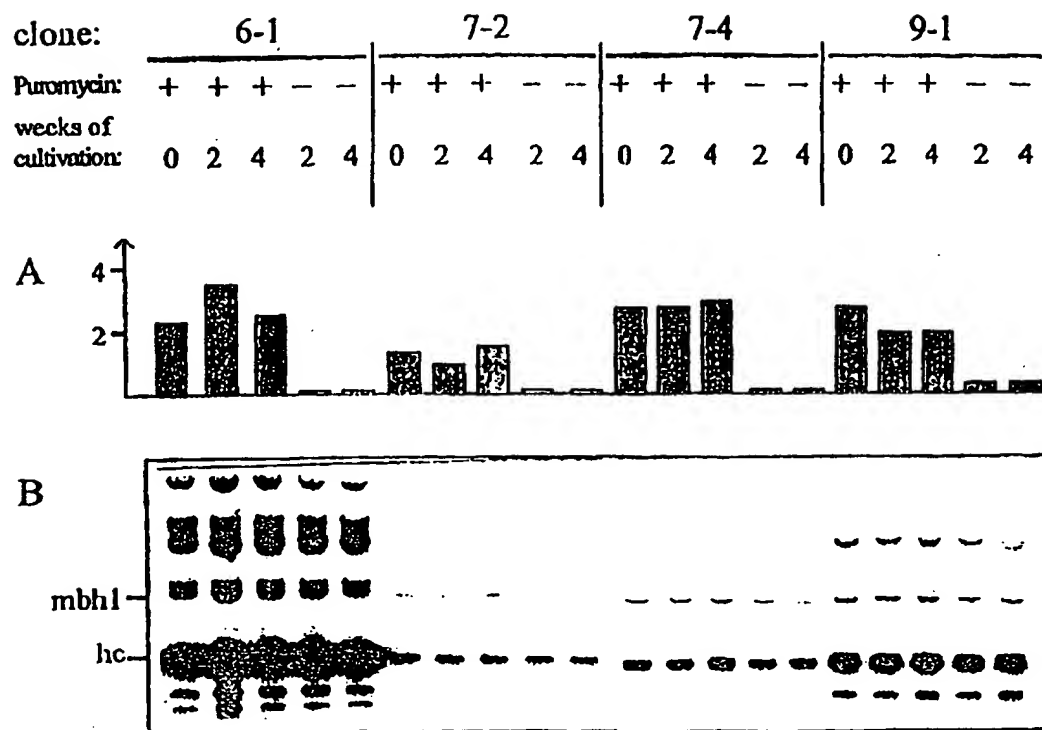


FIG. 10

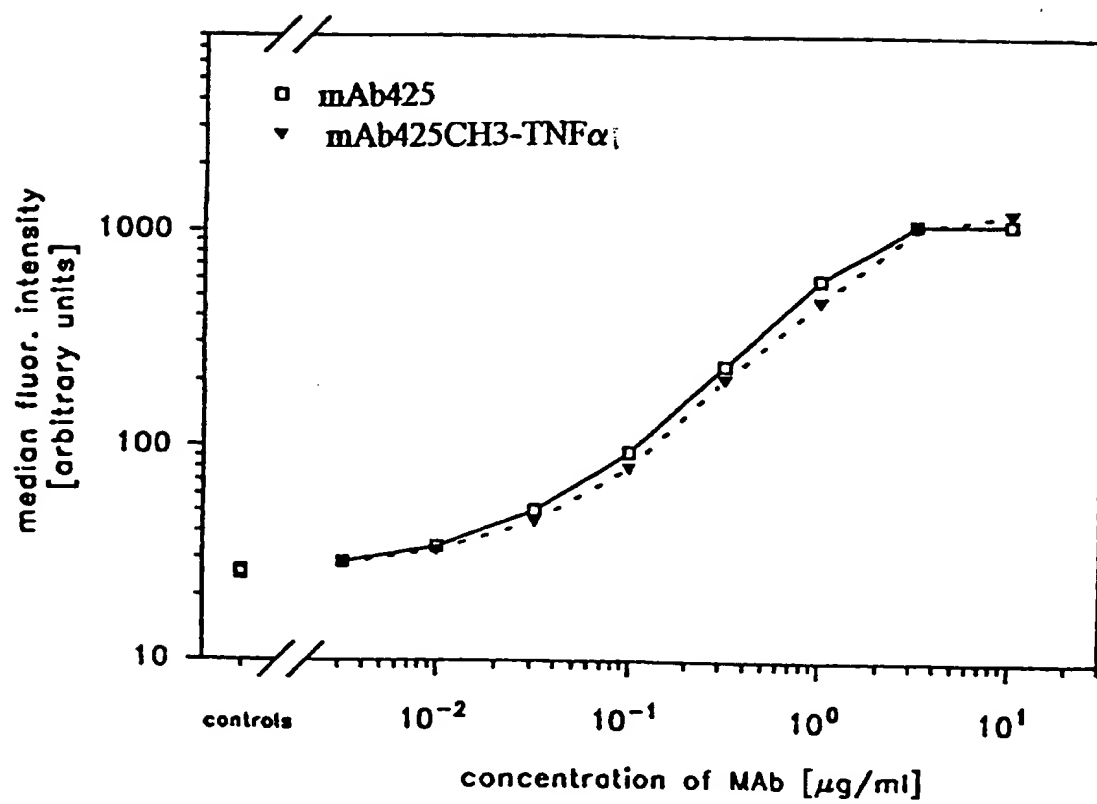


FIG. 11

18/31

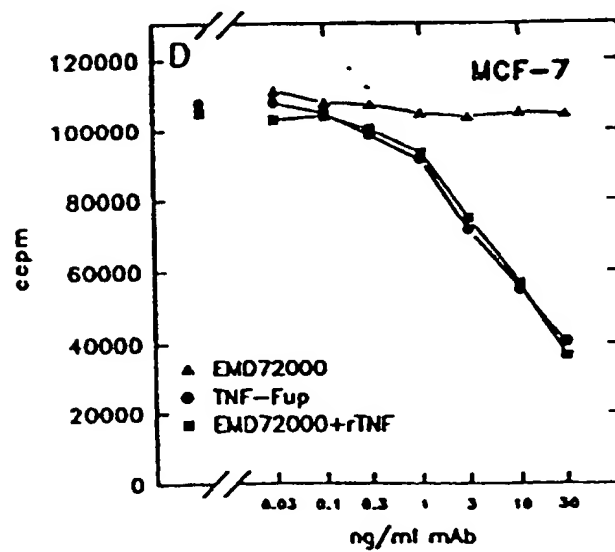
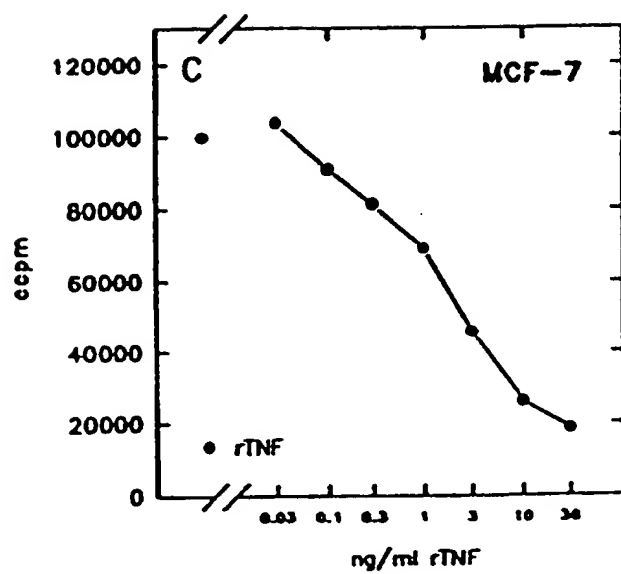


FIG. 12



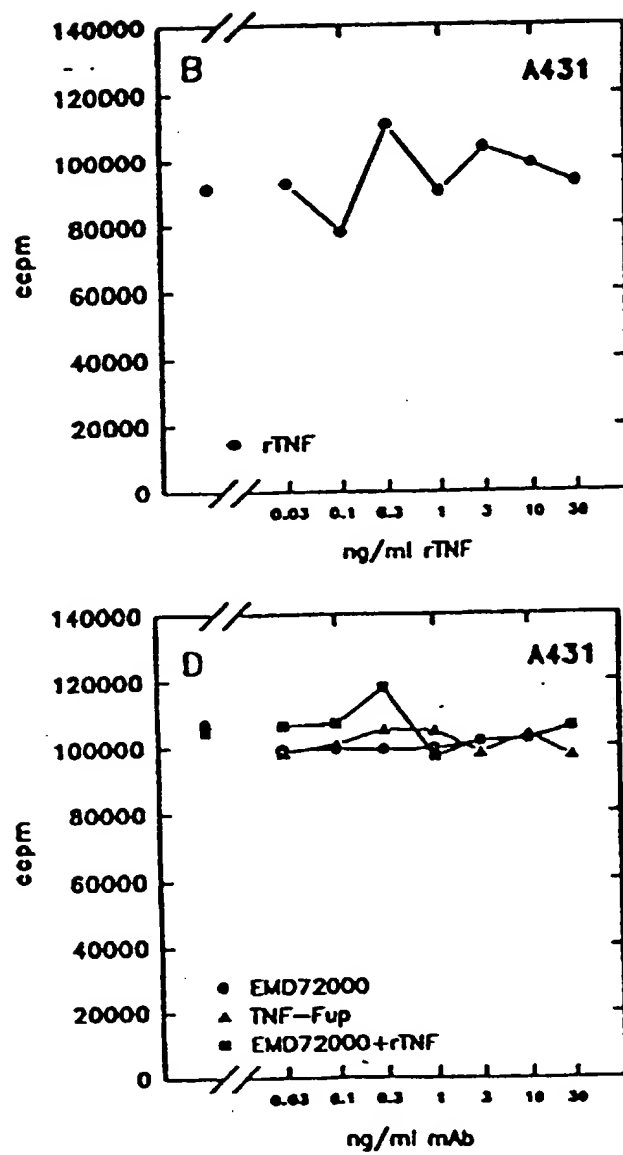


FIG. 13

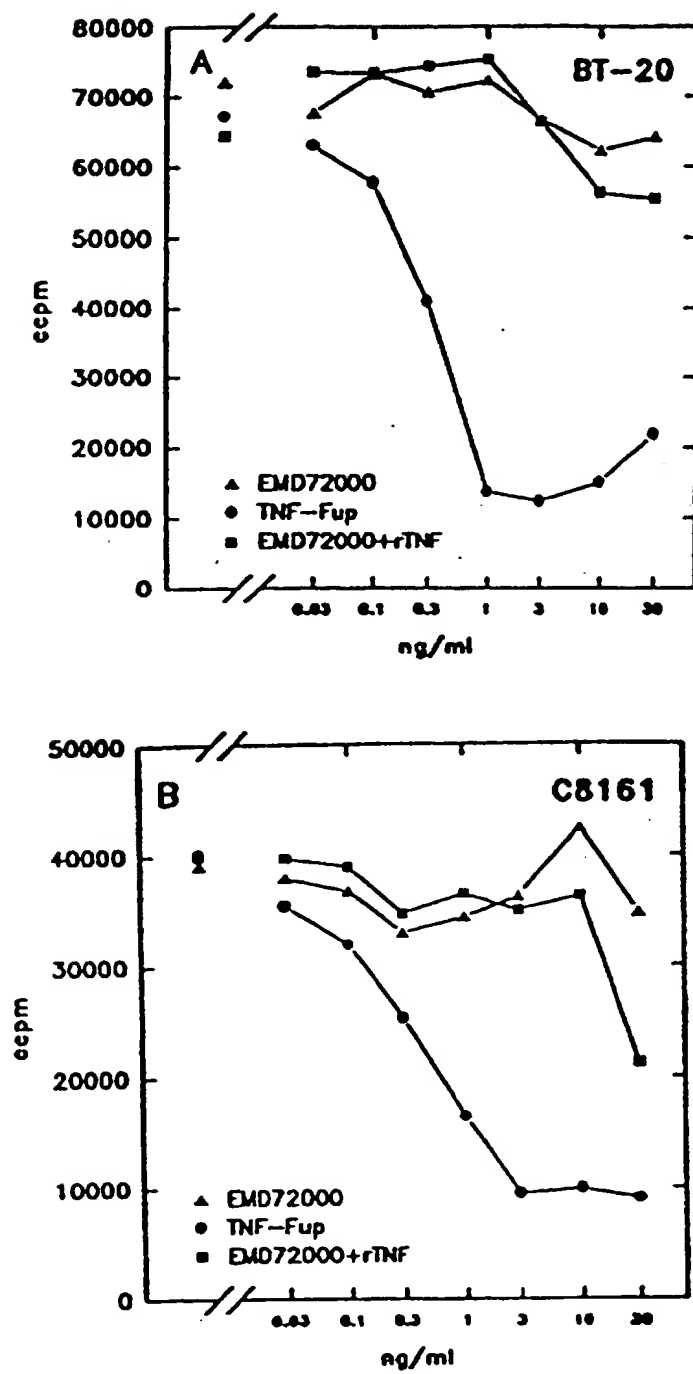


FIG. 14

**Fig.: 15**

TCGATAATGA AAGACCCAC CTGTAGGTTT GGCAAGCTAG CTTAAGTAAC GCCATTTTGC 60  
AAGGCATGGG AAAAATACAT AACTGAGAAT AGAGAAGTTC AGATCAAGGT CAGGAACAGA 120  
GAAACAGGAG AATATGGGCC AAACAGGATA TCTGTGGTAA GCAGTTCCTG CCCCCTCAG 180  
GGCCAAGAAC AGTTGGAACA GGAGAATTGG GCCAAACAGG ATATCTGTGG TAAGCAGTTC 240  
CTGCCCCGCT CAGGGCCAAG AACAGATGGT CCCCAGATGC GGTCCCGCCC TCAGCAGTTT 300  
CTAGACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCCGCCATT 360  
GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA 420  
ATGGGTGGAG TATTTACGGT AAACGCCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC 480  
AAGTACGCCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA 540  
CATGACCTTA TGGGACTTTC CTAATTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC 600  
CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG ACTCACGGGG 660  
ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG 720  
GGACTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG GTAGGCGTGT 780  
ACGGTGGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG 840  
CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCGAGGAACT 900  
GGAAAACCAG AAAGTAACT GGTAAGTTTA GTCTTTTTGT CTTTTATTTC AGGTCCCGGA 960

ATTAAGCTTC GCCACC ATG GGA TGG AGC TGT ATC ATC CTC TTC TTG GTA 1009  
Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val

GCA ACA GCT AC AGGTAAGGGG CTCACAGTAG CAGGCTTGAG GTCTGGACAT 1060  
Ala Thr Ala

ATATATGGGT GACAATGACA TCCAATTGTC CTTTCTCTCC ACAGGT GTC CAC TCC 1115  
Val His Ser

GAC ATC CAG ATG ACC CAG AGC CCA AGC AGC CTG AGC GCC AGC GTG GGT 1163  
Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly

GAC AGA GTG ACC ATC ACC TGT AGT GCC AGC TCA AGT GTA ACT TAC ATG 1211  
Asp Arg Val Thr Ile Thr Cys Ser Ala Ser Ser Ser Val Thr Tyr Met

TAT TGG TAC CAG CAG AAG CCA GGT AAG GCT CCA AAG CTG CTG ATC TAC 1259  
Tyr Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr

GAC ACA TCC AAC CTG GCT TCT GGT GTG CCA AGC AGA TTC AGC GGT AGC 1307  
Asp Thr Ser Asn Leu Ala Ser Gly Val Pr Ser Arg Phe Ser Gly Ser

GGT AGC GGT ACC GAC TAC ACC TTC ACC ATC AGC AGC CTC CAG CCA GAG 1355  
Gly Ser Gly Thr Asp Tyr Thr Phe Thr Ile Ser Ser Leu Gln Pro Glu

GAC ATC GCC ACC TAC TAC TGC CAG CAG TGG AGT AGT CAC ATA TTC ACG 1403  
Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Trp Ser Ser His Ile Phe Thr

TTC GGC CAA GGG ACC AAG GTG GAA ATC AAA CGTGAGTAGA ATTTAAACTT 1453  
Phe Gly Gln Gly Thr Lys Val Glu Ile Lys

TGCTTCCTCA GTTGGATCCA TCTGGGATAA GCATGCTGTT TTCTGTCTGT CCCTAACATG 1513

CCCTGTGATT ATGCGCAAAC AACACACCCA AGGGCAGAAC TTTGTTACTT AAACACCATC 1573

CTGTTTGCTT CTTTCCTCAG GA ACT GTG GCT GCA CCA TCT GTC TTC ATC TTC 1625  
Thr Val Ala Ala Pro Ser Val Phe Ile Phe

CCG CCA TCT GAT GAG CAG TTG AAA TCT GGA ACT GCC TCT GTT GTG TGC 1673  
Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys

CTG CTG AAT AAC TTC TAT CCC AGA GAG GCC AAA GTA CAG TGG AAG GTG 1721  
Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val

GAT AAC GCC CTC CAA TCG GGT AAC TCC CAG GAG AGT GTC ACA GAG CAG 1769  
Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln

GAC AGC AAG GAC AGC ACC TAC AGC CTC AGC AGC ACC CTG ACG CTG AGC 1817  
Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser

AAA GCA GAC TAC GAG AAA CAC AAA GTC TAC GCC TGC GAA GTC ACC CAT 1865  
Lys Ala Asp Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His

CAG GGC CTG AGC TCG CCC GTC ACA AAG AGC TTC AAC AGG GGA GAG TGT 1913  
Gln Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys

TAGAATTCAGCTT TTAACACAGC TCTGGGGTTG TACCCACCCC AGAGGCCAC 1966

GTGGCGGCTA GTACTCCGGT ATTGCGGTAC CCTTGACGC CTGTTTATA CTCCCTTCCC 2026

GTAACCTAGA CGCACAAAAC CAAGTTCAAT AGAAGGGGGT ACAAACCAGT ACCACCACGA 2086

ACAAGCACTT CTGTTTCCCC GGTGATGTCG TATAGACTGC TTGCGTGGTT GAAAGCGACG 2146

GATCCGTTAT CCGCTTATGT ACTTCGAGAA GCCCAGTACC ACCTCGGAAT CTTGATGCG 2206

TTGCGCTCAG CACTCAACCC CAGAGTGTAG CTTAGGCTGA TGAGTCTGGA CATCCCTCAC 2266

CGGTGACGGT GGTCCAGGCT GCGTTGGCGG CCTACCTATG GCTAACGCCA TGGGACGCTA 2326

GTTGTGAACA AGGTGTGAAG AGCCTATTGA GCTACATAAG AATCCTCCGG CCCCTGAATG 2386

CGGCTAATCC CAACCTCGGA GCAGGTGGTC ACAAACCAGT GATTGGCCTG TCGTAACGCG 2446

CAAGTCCGTG GCGGAACCGA CTACTTTGGG TGTCCGTGTT TCCTTTTATT TTATTGTGGC 2506

TGCTTATGGT GACAATCACA GATTGTTATC ATAAAGCGAA TTGGATTGCG GCCGCGAATT 2566

AAGCTTGCCG CCACC ATG GAC TGG ACC TGG CGC GTG TTT TGC CTG CTC GCC 2617

Met Asp Trp Thr Trp Arg Val Phe Cys Leu Leu Ala

GTG GCT CCT GGG GCC CAC AGC CAG GTG CAA CTA GTG CAG TCC GGC GCC 2665

Val Ala Pro Gly Ala His Ser Gln Val Gln Leu Val Gln Ser Gly Ala

GAA GTG AAG AAA CCC GGT GCT TCC GTG AAG GTG AGC TGT AAA GCT AGC 2713

Glu Val Lys Lys Pro Gly Ala Ser Val Lys Val Ser Cys Lys Ala Ser

GGT TAT ACC TTC ACA TCC CAC TGG ATG CAT TGG GTT AGA CAG GCC CCA 2761

Gly Tyr Thr Phe Thr *Ser His Trp Met His* Trp Val Arg Gln Ala Pro

GGC CAA GGG CTC GAG TGG ATT GGC GAG TTC AAC CCT TCA AAT GGC CGG 2809

Gly Gln Gly Leu Glu Trp Ile Gly *Glu Phe Asn Pro Ser Asn Gly Arg*

ACA AAT TAT AAC GAG AAG TTT AAG AGC AAG GCT ACC ATG ACC GTG GAC 2857

*Thr Asn Tyr Asn Glu Lys Phe Lys Ser* Lys Ala Thr Met Thr Val Asp

ACC TCT ACA AAC ACC GCC TAC ATG GAA CTG TCC AGC CTG CGC TCC GAG 2905

Thr Ser Thr Asn Thr Ala Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu

GAC ACT GCA GTC TAC TAC TGC GCC TCA CGG GAT TAC GAT TAC GAT GGC 2953

Asp Thr Ala Val Tyr Tyr Cys Ala Ser *Arg Asp Tyr Asp Tyr Asp Gly*

AGA TAC TTC GAC TAT TGG GGA CAG GGT ACC CTT GTC ACC GTC AGT TCA 3001

*Arg Tyr Phe Asp Tyr* Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser

GGT GAG TGG ATC CTC TGC GCC TGG GCC CAG CTC TGT CCC ACA CCG CGG 3049

Gly Glu Trp Ile Leu Cys Ala Trp Ala Gln Leu Cys Pro Thr Pro Arg

TCA CAT GGC ACC ACC TCT CTT GCA GCC TCC ACC AAG GGC CCA TCG GTC 3097

Ser His Gly Thr Thr Ser Leu Ala Ala Ser Thr Lys Gly Pro Ser Val

TTC CCC CTG GCA CCC TCC TCC AAG AGC ACC TCT GGG GGC ACA GCG GCC 3145

Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala

CTG GGC TGC CTG GTC AAG GAC TAC TTC CCC GAA CCG GTG ACG GTG TCG 3193

Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser

TGG AAC TCA GGC GCC CTG ACC AGC GGC GTG CAC ACC TTC CCG GCT GTC 3241

Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val

CTA CAG TCC TCA GGA CTC TAC TCC CTC AGC AGC GTG GTG ACC GTG CCC 3289

Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro

TCC AGC AGC TTG GGC ACC CAG ACC TAC ATC TGC AAC GTG AAT CAC AAG 3337  
 Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys

CCC AGC AAC ACC AAG GTG GAC AAG AAA GTT GAG CCC AAA TCT TGT GAC 3385  
 Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp

AAA ACT CAC ACA TGC CCA CCG TGC CCA GCA CCT GAA CTC CTG GGG GGA 3433  
 Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly

CCG TCA GTC TTC CTC TTC CCC CCA AAA CCC AAG GAC ACC CTC ATG ATC 3481  
 Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile

TCC CGG ACC CCT GAG GTC ACA TGC GTG GTG GTG GAC GTG AGC CAC GAA 3529  
 Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu

GAC CCT GAG GTC AAG TTC AAC TGG TAC GTG GAC GGC GTG GAG GTG CAT 3577  
 Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His

AAT GCC AAG ACA AAG CCG CGG GAG GAG CAG TAC AAC AGC ACG TAC CGG 3625  
 Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg

GTG GTC AGC GTC CTC ACC GTC CTG CAC CAG GAC TGG CTG AAT GGC AAG 3673  
 Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys

GAG TAC AAG TGC AAG GTC TCC AAC AAA GCC CTC CCA GCC CCC ATC GAG 3721  
 Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu

AAA ACC ATC TCC AAA GCC AAA GGG CAG CCC CGA GAA CCA CAG GTG TAC 3769  
 Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr

ACC CTG CCC CCA TCC CGG GAT GAG CTG ACC AAG AAC CAG GTC AGC CTG 3817  
 Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu

ACC TGC CTG GTC AAA GGC TTC TAT CCC AGC GAC ATC GCC GTG GAG TGG 3865  
 Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp

GAG AGC AAT GGG CAG CCG GAG AAC AAC TAC AAG ACC ACG CCT CCC GTG 3913  
 Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val

CTG GAC TCC GAC GGC TCC TTC TTC CTC TAC AGC AAG CTC ACC GTG GAC 3961  
 Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp

AAG AGC AGG TGG CAG CAG GGG AAC GTC TTC TCA TGC TCC GTG ATG CAT 4009  
 Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His

GAG GCT CTG CAC AAC CAC TAC ACG CAG AAG AGC CTC TCC CTG TCT CCG 4057  
 Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro

GGT AAA ATG GTC AGA TCA TCT TCG CGA ACC CCG AGT GAC AAG CCT GTA 4105  
 Gly Lys Met Val Arg Ser Ser Ser Arg Thr Pro Ser Asp Lys Pro Val

GCC CAT GTT GTA GCA AAC CCT CAA GCT GAG GGG CAA CTG CAG TGG CTG 4153  
 Ala His Val Val Ala Asn Pro Gln Ala Glu Gly Gln Leu Gln Trp Leu

AAC CGC CGG GCC AAT GCC CTC CTG GCC AAT GGC GTC GAG CTG AGA GAT 4201  
Asn Arg Arg Ala Asn Ala Leu Leu Ala Asn Gly Val Glu Leu Arg Asp

AAC CAG CTG GTG GTG CCA TCA GAG GGC CTG TAC CTC ATC TAC TCC CAG 4249  
Asn Gln Leu Val Val Pro Ser Glu Gly Leu Tyr Leu Ile Tyr Ser Gln

GTC CTC TTC AAG GGC CAA GGC TGC CCG TCG ACC CAT GTG CTC CTC ACC 4297  
Val Leu Phe Lys Gly Gln Gly Cys Pro Ser Thr His Val Leu Leu Thr

CAC ACC ATC AGC CGC ATC GCC GTC TCC TAC CAG ACC AAG GTT AAC CTC 4345  
His Thr Ile Ser Arg Ile Ala Val Ser Tyr Gln Thr Lys Val Asn Leu

CTC TCT GCC ATC AAG AGC CCC TGC CAG AGG GAG ACC CCA GAG GGG GCT 4393  
Leu Ser Ala Ile Lys Ser Pro Cys Gln Arg Glu Thr Pro Glu Gly Ala

GAG GCC AAG CCC TGG TAT GAG CCC ATC TAT CTG GGA GGG GTC TTC CAG 4441  
Glu Ala Lys Pro Trp Tyr Glu Pro Ile Tyr Leu Gly Gly Val Phe Gln

CTC GAG AAG GGT GAC CGA CTC AGC GCT GAG ATC AAT CGG CCC GAC TAT 4489  
Leu Glu Lys Gly Asp Arg Leu Ser Ala Glu Ile Asn Arg Pro Asp Tyr

CTC GAC TTT GCC GAG TCC GGA CAG GTC TAC TTT GGG ATC ATT GCC CTG 4537  
Leu Asp Phe Ala Glu Ser Gly Gln Val Tyr Phe Gly Ile Ile Ala Leu

TGATAAGGATCCCCGG GTACCGAGCT CGAATTCAGC TTTTAAAACA GCTCTGGGGT 4593

TGTACCCACC CCAGAGGCC ACCTGGCGGC TAGTACTCCG GTATTGCGGT ACCCTTGATC 4653

GCCTGTTTTA TACTCCCTTC CCGTAACTTA GACGCACAAA ACCAAGTTCA ATAGAAGGGG 4713

GTACAAACCA GTACCACCAC GAACAAGCAC TTCTGTTTCC CCGGTGATGT CGTATAGACT 4773

GCTTGCGTGG TTGAAAGCGA CGGATCCGTT ATCCGCTTAT GTACTTCGAG AAGCCCAGTA 4833

CCACCTCGGA ATCTTCGATG CGTTGCGCTC AGCACTCAAC CCCAGAGTGT AGCTTAGGCT 4893

GATGAGTCTG GACATCCCTC ACCGGTGACG GTGGTCCAGG CTGCGTTGGC GGCCTACCTA 4953

TGGCTAACGC CATGGGACGC TAGTTGTGAA CAAGGTGTGA AGAGCCTATT GAGCTACATA 5013

AGAATCCTCC GGCCCCTGAA TGCGGCTAAT CCAACCTCG GAGCAGGTGG TCACAAACCA 5073

GTGATTGGCC TGTCGTAACG CGCAAGTCCG TGGCGGAACC GACTACTTTG GGTGTCCGTG 5133

TTTCCTTTTA TTTTATTGTG GCTGCTTATG GTGACAATCA CAGATTGTGA TCATAAAGCG 5193

AATTGGATTG CGGCCGGCCG CCACGACCGG TGCCGCCACC ATCCCCTGAC CCACGCCCT 5253

GACCCCTCAC AAGGAGACGA CCTTCC ATG ACC GAG TAC AAG CCC ACG GTG CGC 5306  
 Met Thr Glu Tyr Lys Pro Thr Val Arg

CTC GCC ACC CGC GAC GAC GTC CCC CGG GCC GTA CGC ACC CTC GCC GCC 5354  
 Leu Ala Thr Arg Asp Asp Val Pro Arg Ala Val Arg Thr Leu Ala Ala

GCG TTC GCC GAC TAC CCC GCC ACG CGC CAC ACC GTC GAC CCG GAC CGC 5402  
 Ala Phe Ala Asp Tyr Pro Ala Thr Arg His Thr Val Asp Pro Asp Arg

CAC ATC GAG CGG GTC ACC GAG CTG CAA GAA CTC TTC CTC ACG CGC GTC 5450  
 His Ile Glu Arg Val Thr Glu Leu Gln Glu Leu Phe Leu Thr Arg Val

GGG CTC GAC ATC GGC AAG GTG TGG GTC GCG GAC GAC GGC GCC GCG GTG 5498  
 Gly Leu Asp Ile Gly Lys Val Trp Val Ala Asp Asp Gly Ala Ala Val

GCG GTC TGG ACC ACG CCG GAG AGC GTC GAA GCG GGG GCG GTG TTC GCC 5546  
 Ala Val Trp Thr Thr Pro Glu Ser Val Glu Ala Gly Ala Val Phe Ala

GAG ATC GGC CCG CGC ATG GCC GAG TTG AGC GGT TCC CGG CTG GCC GCG 5594  
 Glu Ile Gly Pro Arg Met Ala Glu Leu Ser Gly Ser Arg Leu Ala Ala

CAG CAA CAG ATG GAA GGC CTC CTG GCG CCG CAC CGG CCC AAG GAG CCC 5642  
 Gln Gln Gln Met Glu Gly Leu Leu Ala Pro His Arg Pro Lys Glu Pro

GCG TGG TTC CTG GCC ACC GTC GGC GTC TCG CCC GAC CAC CAG GGC AAG 5690  
 Ala Trp Phe Leu Ala Thr Val Gly Val Ser Pro Asp His Gln Gly Lys

GGT CTG GGC AGC GCC GTC GTG CTC CCC GGA GTG GAG GCG GCC GAG CGC 5738  
 Gly Leu Gly Ser Ala Val Val Leu Pro Gly Val Glu Ala Ala Glu Arg

GCC GGG GTG CCC GCC TTC CTG GAG ACC TCC GCG CCC CGC AAC CTC CCC 5786  
 Ala Gly Val Pro Ala Phe Leu Glu Thr Ser Ala Pro Arg Asn Leu Pro

TTC TAC GAG CGG CTC GGC TTC ACC GTC ACC GCC GAC GTC GAG TGC CCG 5834  
 Phe Tyr Glu Arg Leu Gly Phe Thr Val Thr Ala Asp Val Glu Cys Pro

AAG GAC CGC GCG ACC TGG TGC ATG ACC CGC AAG CCC GGT GCC TGA 5879  
 Lys Asp Arg Ala Thr Trp Cys Met Thr Arg Lys Pro Gly Ala

CGCCCGCCCC ACGACCCGCA GCGCCCGACC GAAAGGAGCG CACGACCCCA TGAGCTTCGA 5939

TCCAGACATG ATAAGATACA TTGATGAGTT TGGACAAACC ACAACTAGAA TGCAGTGAAA 5999

AAAATGCTTT ATTTGTGAAA TTTGTGATGC TATTGCTTTA TTTGTAACCA TTATAAGCTG 6059

CAATAAACAA GTTAACAACA ACAATTGCAT TCATTTTATG TTTCAGGTTT AGGGGGAGGT 6119

GTGGGAGGTT TTTTAAAGCA AGTAAAACCT CTACAAATGT GGTATGGCTG ATTATGATCC 6179

TGCCTCGCGC GTTTCGGTGA TGACGGTGAA AACCTCTGAC ACATGCAGCT CCCGGAGACG 6239

GTCACAGCTT GTCTGTAAGC GGATGCCGGG AGCAGACAAG CCCGTCAGGG CGCGTCAGCG 6299



GGTGTGGCG GGTGTCGGG CGCAGCCATG ACCCAGTCAC GTAGCGATAG CGGAGTGTAT 6359  
ACTGGCTTAA CTATGCGGCA TCAGAGCAGA TTGTA CTGAG AGTGACCAT ATGTCGGGCC 6419  
GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC CCCCTGACG AGCATCACAA AAATCGACGC 6479  
TCAAGTCAGA GGTGGCGAAA CCCGACAGGA CTATAAAGAT ACCAGGCGTT TCCCCCTGGA 6539  
AGCTCCCTCG TCGCTCTCC TGTTCCGACC CTGCCGCTTA CCGGATACCT GTCCGCCTTT 6599  
CTCCCTTCGG GAAGCGTGGC GCTTTCTCAT AGCTCACGCT GTAGGTATCT CAGTTCGGTG 6659  
TAGGTCGTTT GCTCCAAGCT GGGCTGTGTG CACGAACCCC CCGTTCAGCC CGACCGCTGC 6719  
GCCTTATCCG GTA ACTATCG TCTTGAGTCC AACCCGGTAA GACACGACTT ATCGCCACTG 6779  
GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC 6839  
TTGAAGTGGT GGCCTAACTA CGGCTACACT AGAAGGACAG TATTTGGTAT CTGCGCTCTG 6899  
CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAACCACC 6959  
GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT 7019  
CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC AGTGGAACGA AA ACTCACGT 7079  
TAAGGGATTT TGGTCATGAG ATTATCAAAA AGGATCTTCA CCTAGATCCT TTAAATTA 7139  
AAATGAAGTT TTAAATCAAT CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA 7199  
TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGTCTAT TTCGTTTCAT CATAGTTGCC 7259  
TGA TCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT TACCATCTGG CCCCAGTGCT 7319  
GCAATGATAC CGCGAGACCC ACGCTACCG GCTCCAGATT TATCAGCAAT AAACCAGCCA 7379  
GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT GCAACTTTAT CCGCCTCCAT CCAGTCTATT 7439  
AATTGTTGCC GGGAAGCTAG AGTAAGTAGT TCGCCAGTTA ATAGTGCGCA ACGTTGTTGC 7499  
CATTGCTACA GGCATCGTGG TGTCACGCTC GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG 7559  
TTCCCAACGA TCAAGGCGAG TTACATGATC CCCATGTTG TGCAAAAAAG CGGTTAGCTC 7619  
CTTCGGTCCT CCGATCGTTG TCAGAAGTAA GTTGCCGCA GTGTTATCAC TCATGGTTAT 7679  
GGCAGCACTG CATAATTCTC TACTGTCAT GCCATCCGTA AGATGCTTTT CTGTGACTGG 7739  
TGAGTACTCA ACCAAGTCAT TCTGAGAATA GTGTATGCGG CGACCGAGTT GCTCTTGCCC 7799  
GGCGTCAACA CGGGATAATA CCGCGCCACA TAGCAGAACT TTAAAAGTGC TCATCATTTG 7859  
AAAACGTTCT TCGGGGCGAA AACTCTCAAG GATCTTACCG CTGTTGAGAT CCAGTTTCAT 7919

GTAACCCACT CGTGCACCCA ACTGATCTTC AGCATCTTTT ACTTTCACCA GCGTTTCTGG 7979  
GTGAGCAAAA ACAGGAAGGC AAAATGCCGC AAAAAAGGGA ATAAGGGCGA CACGGAAA 8039  
TTGAATACTC ATACTCTTCC TTTTCAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT 8099  
CATGAGCGGA TACATATTTG AATGTATTTA GAAAAATAAA CAAATAGGGG TTCCGCGCAC 8159  
ATTTCCCCGA AAAGTGCCAC CTGACGTCTA AGAAACCATT ATTATCATGA CATTAACCTA 8219  
TAAAAATAGG CGTATCACGA GGCCCTTTTCG TCTTCAAGAA TTGGTCGATC GACCAATTCT 8279  
CATGTTTGAC AGCTTATCA 8298

FIG. 16

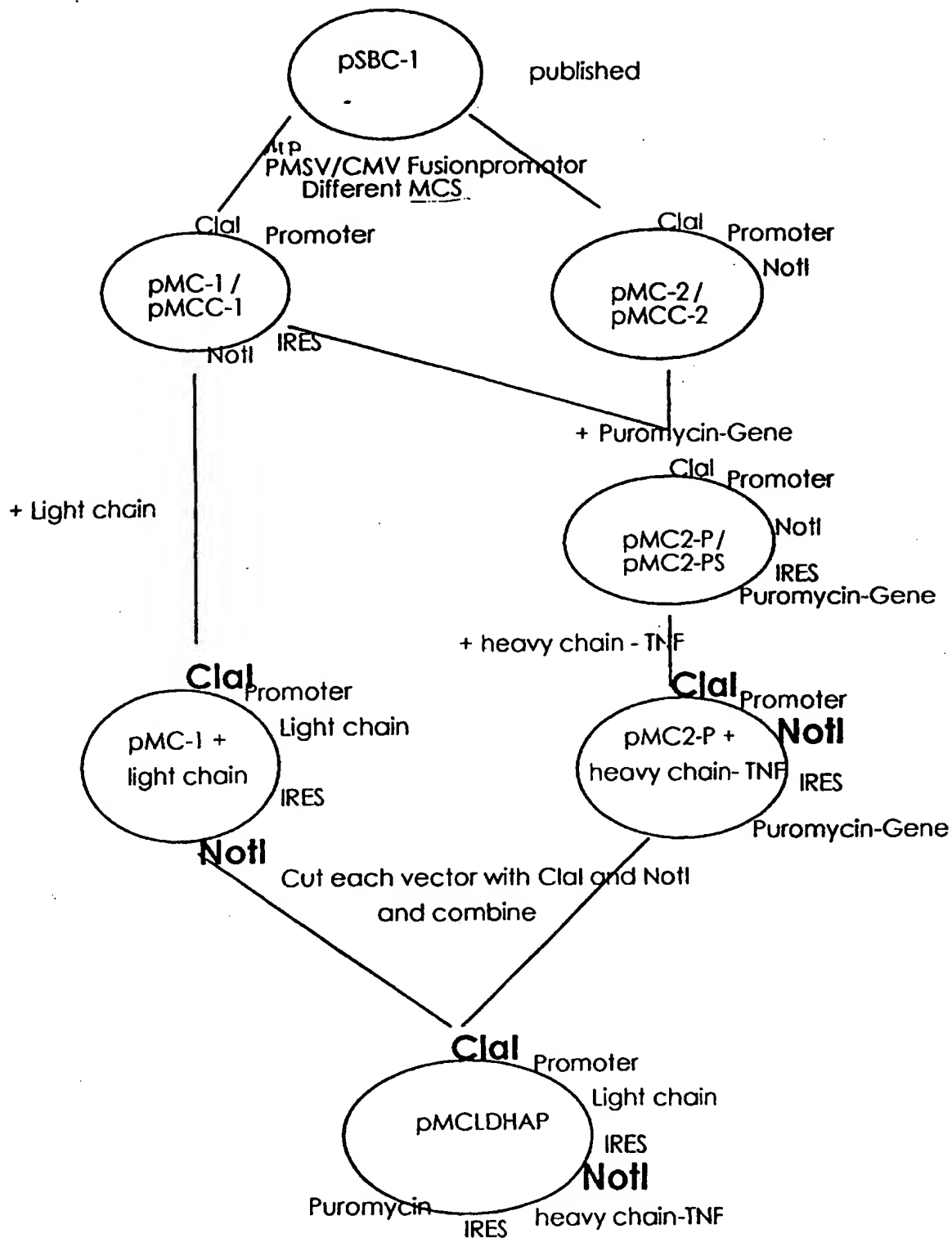


FIG. 17

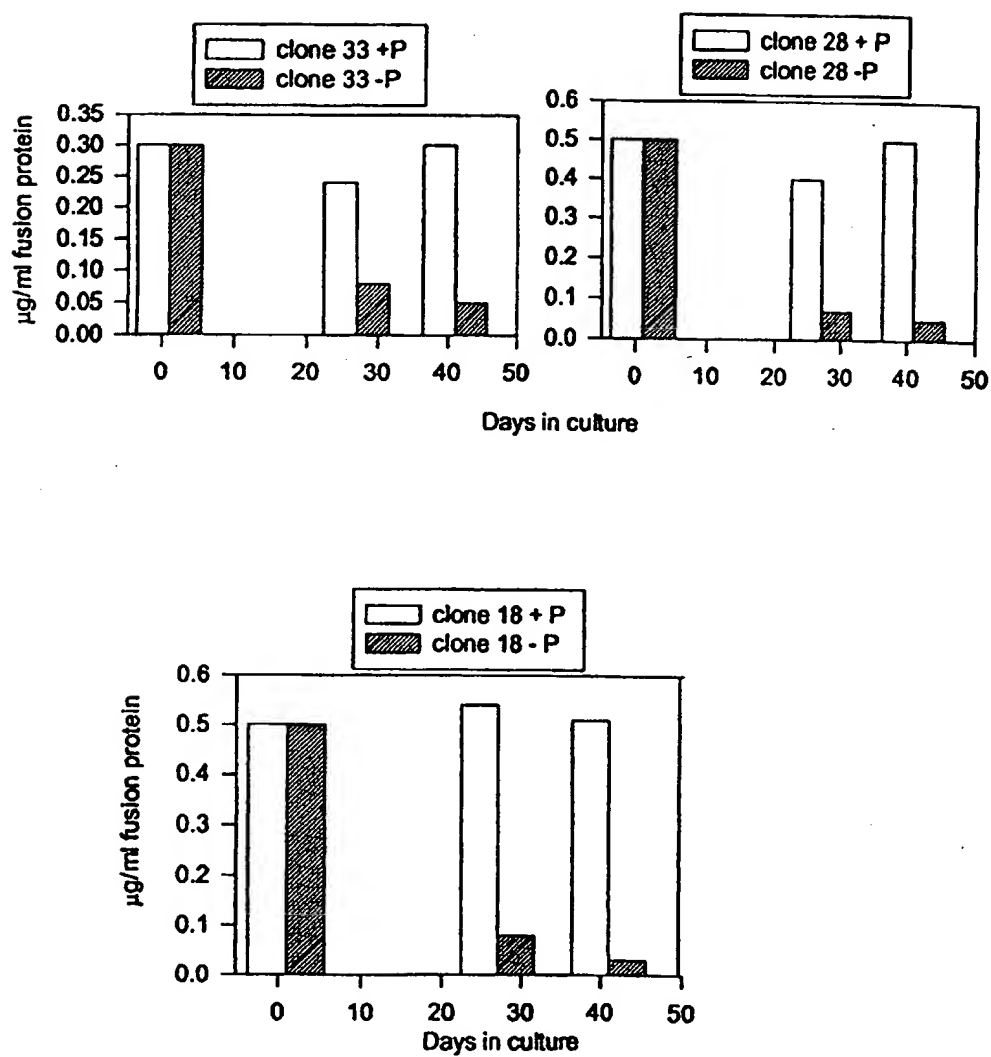
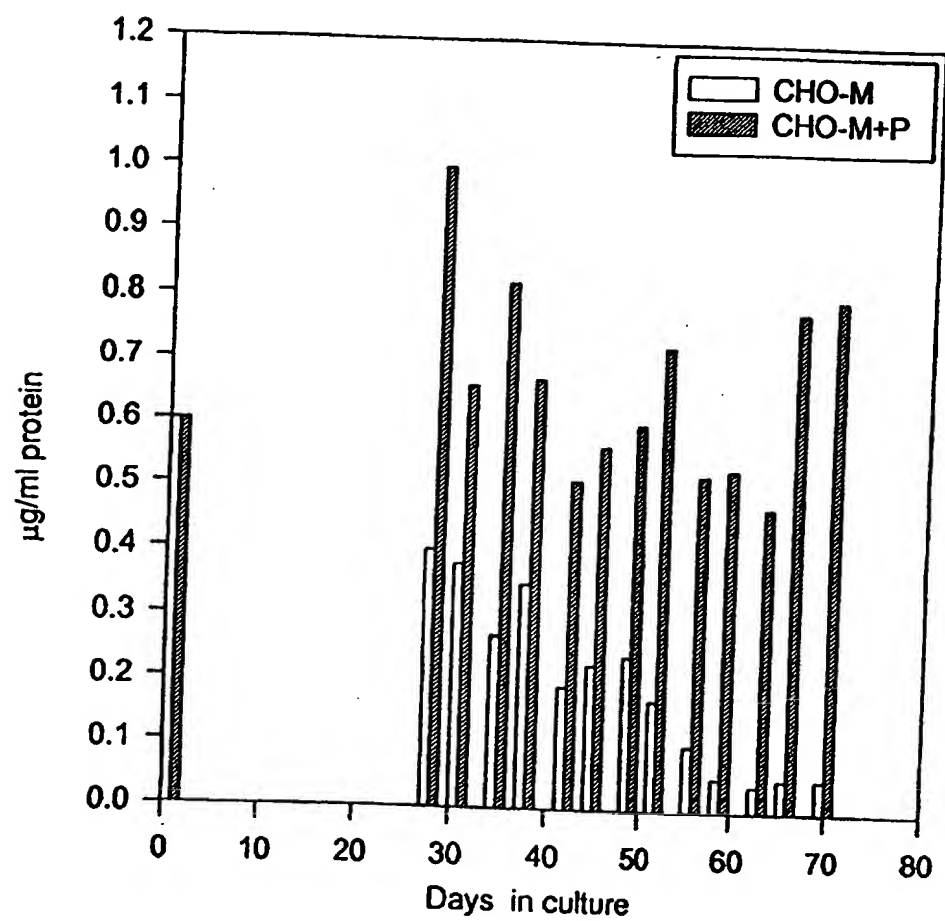


FIG. 18



# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/EP 97/04765

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6	C12N15/85	C12N15/13	C07K16/28	C12N15/43	C12N15/28
	C07K14/525	C12N15/26	C07K14/55	C12N15/19	C07K14/52
	C12N5/10	C12N15/62	C07K19/00	G01N33/53	G01N33/60

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K G01N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 94 24870 A (BIOTRANSPLANT INC ; GEN HOSPITAL CORP (US); LE GUERN CHRISTIAN A (U) 10 November 1994	1, 2, 13, 15
Y	see page 11, paragraph 3 - page 12, paragraph 3 see page 18, paragraph 3 - page 19, paragraph 2; figures 1J-L	3-12, 14, 16
Y	EP 0 659 439 A (MERCK PATENT GMBH) 28 June 1995 cited in the application see abstract see page 3, line 20 - page 4, line 7 see page 6, line 20 - page 9; table 1 see page 11 - page 12; claims see page 13; figure 1	3-12, 14, 16
	-/-	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

9 December 1997

Date of mailing of the international search report

23/01/1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl  
Fax: (+31-70) 340-3016

Authorized officer

Macchia, G

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 97/04765

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>NOEL D. ET AL.: "Analysis of the individual contributions of immunoglobulin heavy and light chains to the binding of antigen using cell transfection and plasmon resonance analysis" JOURNAL OF IMMUNOLOGICAL METHODS, vol. 193, no. 2, 21 June 1996, page 177-187 XP004020811 see page 180; figure 1 see page 182; figure 2</p> <p style="text-align: center;">--- -/--</p>	3

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"Δ" document member of the same patent family

Date of the actual completion of the international search

9 December 1997

Date of mailing of the international search report

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3018

Authorized officer

Macchia, G

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 97/04765

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>CREASEY A.A. ET AL.: "Biological effects of recombinant human tumor necrosis factor and its novel muteins on tumor and normal cell lines"</p> <p>CANCER RESEARCH, vol. 47, 1 January 1987, pages 145-149, XP002049438 see abstract see page 148, right-hand column, paragraph 3 - page 149, left-hand column</p>	5
A	<p>SIDHU R.S. AND BOLLON A.P.: "Tumor necrosis factor analogs: identification of functional domains"</p> <p>ANTICANCER RESEARCH, vol. 9, no. 6, 1989, pages 1569-1576, XP002049439 see page 1569 - page 1570; figure 1 see page 1573, right-hand column - page 1574; figure 4</p>	5
A	<p>WO 92 15683 A (MERCK PATENT GMBH) 17 September 1992 cited in the application see page 10, line 20 - page 13, line 30 see page 23, line 15-20 see page 45, line 28 - page 46, line 12</p>	9-11
A	<p>EVANS M.J. ET AL.: "Rapid expression of an anti-human C5 chimeric Fab utilizing a vector that replicates in COS and 293 cells"</p> <p>JOURNAL OF IMMUNOLOGICAL METHODS, vol. 184, no. 1, 17 July 1995, page 123-138 XP004021009 see abstract see page 127; figure 1</p>	10
A	<p>GROSS G. AND HAUSER H.: "Heterologous expression as a tool for gene identification and analysis"</p> <p>JOURNAL OF BIOTECHNOLOGY, vol. 41, no. 2, 31 July 1995, page 91-110 XP004036927 see page 102, left-hand column, paragraph 3 see page 104, right-hand column - page 105, left-hand column</p>	1,2,12, 14

-/--



# INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 97/04765

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>TAHARA H. ET AL.: "Effective eradication of established murine tumors with IL-12 gene therapy using a polycistronic retroviral vector"</p> <p>THE JOURNAL OF IMMUNOLOGY, vol. 154, no. 12, 15 June 1995, pages 6466-6474, XP002049440 see page 6466 - page 6467 see page 6471, right-hand column - page 6472, left-hand column</p> <p>-----</p>	<p>1,2,13, 15</p>

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 97/04765

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9424870 A	10-11-94	AU 6230394 A CA 2162056 A EP 0706319 A	21-11-94 10-11-94 17-04-96
EP 0659439 A	28-06-95	AU 8159394 A CA 2138928 A CZ 9403306 A HU 70471 A JP 7223968 A NO 944980 A PL 306474 A SK 156294 A ZA 9410282 A	29-06-95 25-06-95 18-10-95 30-10-95 22-08-95 26-06-95 26-06-95 11-07-95 29-08-95
WO 9215683 A	17-09-92	AU 658396 B AU 1340392 A CA 2082160 A CZ 9203327 A EP 0531472 A HU 65687 A MX 9201016 A SK 332792 A US 5558864 A	13-04-95 06-10-92 07-09-92 16-02-94 17-03-93 28-07-94 01-08-93 03-07-96 24-09-96